LES

LANGUES DU MONDE

COLLECTION LINGUISTIQUE

PUBLIÉE PAR

LA SOCIÉTÉ DE LINGUISTIQUE DE PARIS. — XVI

LES LANGUES DU MONDE

PAR

UN GROUPE DE LINGUISTES

SOUS LA DIRECTION DE

A. MEILLET et MARCEL COHEN



PARIS
LIBRAIRIE ANCIENNE ÉDOUARD CHAMPION
5, QUAI MALAQUAIS (6e)

—

1924

COLLECTION LINGUISTIQUE
PUBLIÉE PAR
LA SOCIÉTÉ DE LINGUISTIQUE DE PARIS. — XVI

LES
LANGUES DU MONDE

PAR

UN GROUPE DE LINGUISTES

SOUS LA DIRECTION DE

A. MEILLET ET Marcel COHEN



PARIS
LIBRAIRIE ANCIENNE ÉDOUARD CHAMPION
5, QUAI MALAQUAIS (6e)

—

1924

AVANT-PROPOS

Les auteurs du présent recueil ont voulu donner une idée de la répartition des langues dans le monde, en tenant compte de l'histoire de ces langues.

Un ouvrage collectif tel que celui-ci manque nécessairement d'unité.

Les divers domaines sont inégalement étudiés : depuis le groupe indo-européen, pour lequel il existe une grammaire comparée minutieusement établie, jusqu'aux langues américaines pour lesquelles on n'a d'ordinaire pas dépassé, et souvent pas atteint, le stade de la description, il se rencontre toutes sortes de cas.

D'autre part, les auteurs ont des formations, des habitudes d'esprit différentes.

Il aurait été vain de chercher à obtenir des exposés semblables. En imposant un même plan, une même manière de traiter le sujet, on n'aurait abouti à donner que l'illusion du parallélisme. L'essentiel était de marquer, pour chaque domaine, l'état actuel des connaissances. Les auteurs n'y pouvaient parvenir qu'à la condition d'avoir leur liberté et de conformer leur exposé aux conditions propres de leur discipline spéciale.

Les indications données sur les limites des langues et sur la statistique ne visent pas à la précision ; il n'est

utile de donner des nombres précis que là où le sujet admet la précision, et où les recherches ont été faites avec exactitude. Or, il y a beaucoup de domaines linguistiques troubles, de régions où il se parle plusieurs langues, ou bien où les dialectes n'ont pas de limites arrêtées. Le degré d'unité linguistique varie d'un pays à l'autre. Deux sujets parlant l'arabe, l'un du Maroc, l'autre de Syrie, s'entendent plus ou moins ; mais, en employant chacun leur parler, ils ne sauraient guère se comprendre qu'après accommodation ; pourtant l'un et l'autre sont dits de langue arabe. Dans toute la France on comprend le français ; mais un paysan provençal qui emploie dans sa famille le parler local et qui ne parle le français qu'avec un fort « accent », est-il de langue française au même titre qu'un bourgeois de Paris qui emploie uniquement le français commun ? La variété des cas exclut toute statistique rigoureuse ; et, même dans la mesure où une statistique serait possible, elle n'a été faite nulle part d'une manière précise, nuancée — et désintéressée. Seule, une enquête dirigée par des linguistes sans parti pris national permettrait une statistique véritable, et la difficulté en serait extrême. Il n'y a pas une région dont l'état linguistique ait été décrit d'une manière complète.

Aucun lecteur ne se posera sans doute la question naïve du nombre de langues ou de familles de langues existant dans le monde.

Il y a autant de parlers distincts qu'il y a de localités et, dans ces localités, de groupes ayant une autonomie. Il n'y a qu'un français littéraire, mais le français varie d'une localité du domaine français à l'autre.

Quant aux familles de langues, le nombre s'en réduit au fur et à mesure que les recherches s'approfondissent. Et il suffit d'avoir compris la définition même du terme « famille de langues » pour apercevoir aussitôt qu'une famille n'est jamais comparable à aucune autre. Une famille de langues se définit, on le sait, par une langue commune initiale ; mais ces langues communes ne sont pas contemporaines les unes des autres : le bantou commun est assurément beaucoup plus récent que l'indo-européen commun, lequel est, suivant toute apparence, plus récent que le sémitique commun.

Les précisions de la linguistique ne sont pas de celles qui s'énoncent par des nombres.

Les noms de certains idiomes dont il n'y avait rien à dire, sinon qu'ils font partie de tel ou tel ensemble, figurent seulement à l'index, celui-ci visant à être complet — sans qu'on se flatte d'avoir approché de cet idéal.

La variété des sujets — et des auteurs — a empêché d'unifier le mode de références bibliographiques. Chaque partie est à cet égard indépendante.

L'index qui indique toutes les mentions de chaque idiome a permis de supprimer à l'intérieur du volume les références d'un chapitre à l'autre.

Les cartes qui figurent ici sont une nouveauté dans un ouvrage sur l'ensemble des langues. La nécessité en était évidente. Elles ne sont pas aussi satisfaisantes qu'on aurait pu le désirer. Ce n'est du reste que dans un nombre de cas relativement restreint qu'on aurait pu marquer exactement la coexistence de plusieurs langues sur un même

domaine, la densité de la population parlant chaque langue, le degré d'unité de chaque idiome, etc. Il y a de ce côté un progrès immense à faire.

Outre les linguistes français qui ont bien voulu délaisser leurs occupations propres pour se rendre à mon appel, je tiens à remercier deux savants russes, M. Elisséèv et le prince N. Troubetzkoy, dont la collaboration a été précieuse.

M. Marcel Cohen a, dès le début, accepté d'être le secrétaire de la rédaction. Il a rempli ses fonctions avec tout le zèle possible, unifiant l'aspect du livre autant qu'il était désirable, revoyant les épreuves et surveillant tout. Sans lui, le livre n'existerait pas.

Il a fallu à l'éditeur Champion beaucoup de courage pour entreprendre une publication aussi coûteuse. Il a déjà rendu trop de services à la linguistique pour qu'on soit surpris de cette nouvelle preuve de son dévouement à la science. L'imprimerie Protat s'est acquittée d'un travail difficile avec sa maîtrise ordinaire. Nous remercions l'éditeur et l'imprimeur ; nous espérons que le public leur saura gré de leur effort.

A. MEILLET.

TABLEAU GÉNÉRAL

DE

TRANSCRIPTION ET DE NOTATION PHONÉTIQUE

Les mots imprimés en italiques sont en principe en écriture phonétique suivant le système adopté pour l'ouvrage.

Dans ce système, les caractères dont il n'est rien dit ci-dessous sont à lire comme en français.

Certains signes complémentaires sont ajoutés et définis dans quelques-uns des chapitres.

SÉRIES CONSONANTIQUES

Aspiration	est marquée par	ʻ après le signe	*tʻ*
Mouillure	— —	′ —	*t′*
Prononciation cacuminale	— —	un point au-dessous	*ṭ*
Emphase	— —	un trait au-dessous	*ṯ*
Affrication	— —	l'addition d'une occlusive et d'une spirante	*tš*, *mp*
Sourdité des consonnes douces et des sonantes	— —	l'emploi du signe de la sonore correspondante en petites capitales : ᴅ, ɴ	
Indétermination d'une consonne	— —	l'emploi de la romaine au milieu des italiques	*a*p*a*
Vocalisation (emploi d'une consonne ou d'une sonante comme voyelle)	— —	un rond au-dessous	*n̥*

REMARQUE. — Exceptionnellement, les consonnes grecques θ φ χ sont transcrites *th*, *ph*, *kh*.

Les mêmes consonnes grecques servent pour la translitération de l'étrusque. (Pour la même langue la consonne *c* est reproduite telle quelle.)

Les transcriptions, usuelles en linguistique, de certaines affriquées par un seul caractère n'ont pas été employées ; donc dans le présent ouvrage on trouvera : *tš* = *č* ; *dž* = *ǵ* (*j*), *ts* = *c*, etc.

CONSONNES PARTICULIÈRES

Spirantes	*sourdes*	*sonores*
Laryngales	*ḥ* (ح de l'arabe) Sorte d'*h* fort prononcé en position de voix chuchotée	ʿ (ع de l'arabe) Émission de voix avec constriction glottale.
Vélaires	*ḫ* (*ch* allemand prononcé à l'alsacienne) *x* (*ch* allemand ordinaire dans « ach »)	*g* son de l'arabe غ, de *g* allemand du Nord entre deux voyelles, de γ grec moderne, etc.
Prépalatales	*š* (*ch* du français) *x'* (*ch* de l'allemand dans « ich »)	*ž* (j du français)
Interdentales	*þ* (*th* anglais fort)	*d* (*th* anglais faible).
Bilabiales	*p̄* (*p* spirant)	*ƀ* (*b* spirant)

REMARQUE. — *s* est toujours sourd.

LIQUIDES ET SEMI-VOYELLES

Vélaires et postpalatales. — *ṅ* comme *n* dans anglais « sing ».
ł (*l* profond du russe).
ρ (*r* du gosier).

Prépalatales. — *ñ* comme *gn* du français.
y comme dans « yeux ».

Labiales. — *w* comme *ou* dans « oui ».
ẅ comme *u* dans « puits ».

REMARQUE. — *r* est toujours lingual.

OCCLUSIVE GLOTTALE ET APOSTROPHE

L'apostrophe note exclusivement l'occlusive glottale ou coup de glotte '.

Les élisions sont indiquées par un tiret, ou ne sont pas signalées (exemple : *l arbre* ou *l-arbre*, non « l'arbre »).

SÉRIES VOCALIQUES

Quantité.	Longues *ā* brèves	*ă*
Fermeture.	Point au-dessous	*ọ*
Ouverture.	Crochet —	*ǫ*
Nasalité.	Tilde au-dessus	*ã*
Accent principal :		*á*
— secondaire :		*à*
— modulé :		*ȃ*

TIMBRES PARTICULIERS

u = *ou* du français
ü = *u* —
ö = *e* muet — (français « pr*e*nez »)
ä = voyelle entre *a* et *ę*
å = — — *a* et *o*
į = *i* large (*y* du slave)
ə = voyelle non déterminable.

ABRÉVIATIONS

Seules ont été employées les abréviations usuelles de certains termes grammaticaux.

Il y a lieu de noter en outre :
MSL = Mémoires de la Société de linguistique de Paris.

LISTE CHRONOLOGIQUE

DES

PRINCIPAUX OUVRAGES CONTENANT DES TABLEAUX DE FAMILLES DE LANGUES

HERVÁS (el Abate don Lorenzo). — Catálogo de las lenguas de las naciones conocidas, y numeración, división, y clases de estas según la diversidad de sus idiomas y dialectos. Madrid, 6 volumes, de 1800 à 1805.

MITHRIDATES, oder Allgemeine Sprachenkunde, mit dem Vater Unser als Sprachprobe in bey nahe fünfhundert Sprachen und Mundarten. Volume I (J. C. Adelung), Berlin, 1806 ; volume II (J. C. Adelung et J. S. Vater), 1809 ; volume III (J. C. Adelung et J. S. Vater), 1812-1816 ; volume IV (J. S. Vater, Fr. Adelung, W. von Humboldt), 1817.

(Cet ouvrage donne une riche bibliographie des travaux antérieurs, en particulier HERVAS, Idea dell' Universo, volumes XVII à XXI.)

BALBI (Adrien). — Atlas ethnographique du globe, ou classification des peuples anciens et modernes d'après leurs langues, Paris, 1826 (Tableaux sans cartes).

AUER. — Sprachhalle, oratio dominica oder das Vater Unser in DCCLXII Sprachen, Vienne, 1844-1847.

VATER (J. S.). — Literatur der Grammatiken, Lexica und Wœrtersammlungen aller Sprachen der Erde, 2e éd., par B. Jülg, Berlin, 1847.

LEPSIUS (C. R.) (1re édition de l'Alphabet universel de transcription : Berlin, 1855). — Standard alphabet for reducing unwritten languages and foreign graphic systems to a uniform orthography in european letters, Londres-Berlin, 1863 (Contient une classification généalogique des langues).

MAX MÜLLER (F.). — Lectures on the Science of language, Londres, 1861 — [The Science of language Londres, 1891.]

WHITNEY (W. D.). — Language and the Study of language, Londres, 1867. [*En français*, La vie du langage, Paris, 1875.]

Reise der œsterreichischen Fregatte *Novara* um die Erde in den Jahren 1857, 1858, 1859. Linguistischer Theil, von Dr. Friedrich Müller, Vienne, 1867.

Müller (Friedrich). — Grundriss der Sprachwissenschaft, 4 tomes, Vienne, 1876-1877, 1882, 1884-1887, 1888. (Le seul grand ouvrage descriptif depuis *Mithridates*, grammaire résumée et texte spécimen d'une langue de chaque groupe).

Hovelacque (A.). — La linguistique (Histoire naturelle du langage). Paris, 1re édition 1876, dernière réimpression, 1921.

Sayce (A. H.). — Introduction to the Science of language, 2 vol., Londres 1879.

Finck (F. N.). — Die Sprachstämme des Erdkreises, Leipzig, 1909 (Premier essai depuis *Mithridates* de tableau complet de toutes les langues connues, rangées par familles, avec liste-index).

— Die Haupttypen des Sprachbaus, Leipzig, 1909 (description de 8 langues de type différent).

Les ouvrages concernant une seule partie du monde sont indiqués dans les bibliographies particulières à l'intérieur du livre (ainsi voir pour l'Afrique, p. 478, pour l'Amérique, p. 605).

Une abondante liste de périodiques est donnée p. 311-312.

Pour compléter la bibliographie sur presque tous les domaines, consulter les recueils suivants :

Literatur-Blatt für orientalische Philologie, éditeurs : E. Kuhn et J. Klatt. Vol. I-IV, 1883-1887.

Orientalische Bibliographie, éditeurs : A. Müller, puis E. Kuhn, puis E. Kuhn et L. Schermann, puis L. Schermann. Vol. I-XXIV, Berlin, 1888-1914. (Le dernier fascicule paru contient la bibliographie pour 1909-1910.)

Rivista degli studi Orientali, Rome. — Bollettini bibliografici, dans : volume VI, 1914-1915 ; volume VII, 1916-1918 ; volume VIII, 1919.

INTRODUCTION

La trop fameuse classification en langues isolantes, agglutinantes et flexionnelles ne se laisse pas poursuivre exactement, et, pour autant qu'elle se laisse formuler, elle n'a ni portée scientifique ni utilité pratique.

La seule classification linguistique qui ait une valeur et une utilité est la classification généalogique, fondée sur l'histoire des langues. C'est celle qui fournira le plan du présent ouvrage.

Le principe en est connu : lorsqu'une langue est parlée sur un domaine étendu et que les individus qui l'emploient viennent à cesser d'avoir les relations régulières et continues qui maintenaient l'unité de langue, les changements qui ont lieu dans les diverses parties du domaine ne sont pas identiques ; et, au bout d'un temps, variable suivant les cas, les différences entre les parlers locaux qui continuent la langue d'abord commune deviennent telles que les occupants des diverses régions cessent de se comprendre aisément entre eux. Dans la mesure où les habitants de provinces différentes cessent de se comprendre, on peut dire que la langue commune est remplacée par des langues nouvelles.

Ainsi le latin, qui était la langue commune de la partie occidentale de l'Empire romain, s'est développé de manières différentes dans les diverses parties de l'Italie, de la péninsule hispanique, de la Gaule, etc., à partir du moment où la civilisation antique s'est disloquée. Et, sans qu'ils se soient à aucun moment rendu compte du fait qu'ils ne parlaient plus latin, les habitants des diverses régions se sont trouvés employer des langues nettement distinctes les unes des autres — et toutes très

différentes du latin — au IXe siècle après J.-C. Les parlers italiens, hispaniques, gallo-romans, rhéto-romans, roumains sont tous du latin, mais du latin qui a évolué de manières diverses.

Les évolutions divergentes subies par le latin ont eu lieu à date historique. Des évolutions divergentes du même type ont eu lieu à des moments sur lesquels on n'a aucun témoignage historique. Mais les faits linguistiques suffisent souvent à l'indiquer. Par exemple, les concordances qu'on observe entre le sanskrit, la langue de l'Avesta, le grec ancien, le latin ancien, le vieil irlandais, le gotique, le vieux slave, l'ancien arménien, etc., ne peuvent s'expliquer si ces diverses langues ne résultent pas d'évolutions diverses d'une seule et même langue, dite indo-européenne.

De là résulte la définition des termes : *famille de langues* et *langues parentes*.

Une famille de langues est l'ensemble des parlers plus ou moins différenciés entre eux qui continuent une même langue commune. Ainsi la famille *romane* ou *néo-latine* est l'ensemble des langues qui sont des formes diversement évoluées du latin. La famille *indo-européenne* est l'ensemble des langues qui sont des formes diversement évoluées de l'indo-européen.

Des parlers *parents*, des langues *parentes* sont des parlers, des langues qui font partie d'une famille ainsi définie.

Cette définition est purement historique. Elle n'implique entre les langues considérées aucune concordance ni de type général ni de détail. Il subsiste d'ordinaire des concordances, et souvent de très considérables. Mais on n'en peut jamais prévoir l'étendue ni le caractère particulier.

Des langues parentes peuvent différer entre elles au point que la parenté ne soit plus reconnaissable à aucun trait. Par exemple, entre le polonais et l'anglais actuels, les ressemblances sont très petites et dans la structure d'ensemble et dans le détail. On sait néanmoins que l'anglais et le polonais appartiennent à un même groupe : le groupe indo-européen ; mais c'est parce que l'on peut établir, par des procédés indirects, que l'un et l'autre résultent d'une série de transformations qui ont eu lieu de manière

continue entre l'indo-européen commun et l'anglais d'une part, le polonais de l'autre.

A l'intérieur d'une même famille, il peut se constituer des familles nouvelles. Une langue qui est une forme évoluée d'une langue antérieure peut se différencier à son tour en langues diverses. De même que l'indo-européen s'est différencié en indo-iranien, grec, latin, germanique, etc., le latin s'est différencié en italien, espagnol, portugais, provençal, français, etc. ; le germanique s'est différencié en gotique, scandinave, allemand, anglais, etc.

Dès lors, il est impossible d'établir que deux langues ne sont pas parentes. Il est aisé de montrer que le flamand n'a aucun des caractères auxquels on reconnaît une langue romane. Mais le flamand est une langue germanique, donc l'une des langues indo-européennes ; il est donc indo-européen comme les langues romanes, qui sont du latin évolué. La parenté n'est pas la parenté proche des langues romanes ; c'est la parenté indo-européenne, plus lointaine. Dès lors, le fait que le finno-ougrien et l'indo-européen sont très différents ne prouve pas que les deux groupes ne continuent pas un original commun, beaucoup plus éloigné dans le passé.

Toute classification généalogique des langues exprime seulement le fait historique d'une communauté linguistique ayant existé à un moment du passé. On peut donc toujours imaginer que, en montant plus haut dans le passé, on découvrirait une communauté linguistique à laquelle se rattacheraient deux langues ou deux groupes de langues qui ne se laissent pas expliquer par les communautés linguistiques actuellement déterminées.

Le difficile est de faire la preuve.

Toutes choses égales d'ailleurs, il y a chance pour que deux langues parentes divergent d'autant plus qu'elles sont séparées depuis plus longtemps. Par suite, la preuve d'une parenté de langues est plus difficile à administrer au fur et à mesure que la communauté supposée est plus loin de l'époque où sont attestées les langues à considérer.

Il est facile de prouver que les langues romanes sont parentes entre elles, parce que, entre la période où le latin a commencé à se briser en parlers divers et les plus anciens textes des langues romanes, il ne s'est écoulé que peu de siècles. Et, si l'on est bien éclairé sur la parenté des langues indo-européennes entre elles, c'est que l'on a, sur plusieurs de ces langues, des données de plusieurs siècles antérieures à l'époque chrétienne et, pour presque toutes, des données antérieures au x[e] siècle après J.-C. Si l'on devait faire la théorie des langues indo-européennes seulement avec les formes actuelles de ces langues, on en apercevrait la parenté ; mais il serait impossible d'en construire une grammaire comparée précise comme on le fait grâce aux anciens monuments qu'on possède de diverses langues.

Or, pour la plus grande partie des langues du monde, les données dont on dispose sont le plus souvent ou peu anciennes ou seulement modernes. Ces données permettent de déterminer des communautés qui remontent à un nombre restreint de siècles dans le passé. Mais, là où il faudrait admettre des communautés de beaucoup antérieures à la communauté indo-européenne ou à la communauté sémitique, les changements intervenus entre ces périodes, qu'on devrait supposer antérieures de plusieurs millénaires à l'ère chrétienne, et l'époque actuelle sont tellement profonds que toute démonstration devient illusoire.

Du reste, la classification généalogique a d'autant moins de valeur, soit pour permettre l'étude historique des langues, soit pour en faciliter l'apprentissage, que les langues considérées conservent moins de traits communs. Si, ne connaissant de l'anglais, du français, du polonais et de l'arménien que les formes actuelles, on devait faire l'histoire de ces langues à l'aide de la seule comparaison, on n'irait évidemment pas loin ; et, de savoir que ces quatre langues sont des formes diverses prises par un même idiome, parlé il y a quatre ou cinq mille ans, n'aiderait ni à les parler ni à en comprendre le développement. Intéressante pour l'historien, cette conclusion ne serait guère utile directement au linguiste.

Si le fait que certaines langues sont des formes prises avec le

temps par une langue commune est une donnée précieuse pour l'histoire, il n'en faut pas, même à ce point de vue, exagérer la portée.

Dans toute langue, il y a lieu de considérer, à côté du fonds transmis au cours des générations, les forces — pour la plupart inconnues — qui ont déterminé les changements. Le français est du latin transformé. Mais la transformation a été radicale : par la structure générale comme par le détail, le français est très différent du latin ; par suite, les forces d'où vient la transformation entrent dans l'état actuel du français pour une part qui ne se laisse ni mesurer ni même évaluer approximativement, mais qui dépasse en importance la part de l'élément latin ancien. Les linguistes insistent sur l'état de choses initial, en l'espèce sur le latin, parce qu'ils le connaissent plus ou moins, soit par des témoignages directs — c'est le cas du latin —, soit par des procédés comparatifs ; ils ne parlent guère des forces de transformation, parce que jusqu'ici on n'est pas arrivé à se former d'idées précises à ce sujet. Mais il ne faut pas se laisser abuser par l'imperfection actuelle des connaissances.

D'ailleurs, il y a une part de la langue — et souvent une très large part — dont la classification généalogique ne tient pas compte. Ce sont les *emprunts*. Qui voudrait expliquer l'ensemble du vocabulaire français actuel par le développement du « latin vulgaire » d'époque impériale, se trouverait devant un reste énorme dont il ne pourrait rien faire : la plus grande partie du vocabulaire français a été « empruntée » au latin des livres, ancien ou médiéval, et à des langues diverses : allemand, arabe, italien, espagnol, anglais, etc. Toutes les fois que, par suite d'altérations internes de la langue ou du besoin d'exprimer des notions nouvelles, il a fallu au français des mots nouveaux, ces mots ont été pris systématiquement au latin écrit, et adaptés pour la forme et pour le sens ; ainsi le latin écrit, ancien et médiéval, entre dans la composition du vocabulaire français moderne pour autant ou pour plus que le « roman commun » (le latin vulgaire) des romanistes. A une époque plus récente, le vocabulaire grec a été largement mis à contribution. La structure des phrases a

aussi subi des influences étrangères, surtout celle du latin écrit.

L'« emprunt » est particulièrement visible là où il est fait à une langue de type tout différent ; ainsi le persan actuel, qui est une langue indo-européenne, doit à peu près tout son vocabulaire des choses de l'esprit à l'arabe, qui est une langue sémitique.

*
* *

La part de la langue qui se maintient de génération en génération en se transformant progressivement est le système grammatical. Sans doute la morphologie latine est loin de la morphologie indo-européenne, et la morphologie française, loin du système latin ; il y a là trois structures distinctes. Mais beaucoup de particularités du système latin se retrouvent en français, et beaucoup de particularités du système indo-européen se retrouvent en latin. Et de plus, on voit comment le système français est sorti du système latin, et le système latin du système indo-européen. Il y a une continuité dans la morphologie, et c'est cette continuité qui permet le classement. Le classement généalogique est fondé sur la continuité de la morphologie.

On reconnaît la parenté à des détails comme la concordance de *il est, ils sont*, en français, *er ist, sie sind*, en allemand, beaucoup plus qu'à des ressemblances de structure générale. C'est donc avec les formes « fortes » et surtout avec les formes anomales des langues indo-européennes attestées qu'on parvient à retrouver la norme de l'indo-européen, et par suite, à poser la grammaire comparée de l'ensemble du groupe. Et en effet, du fait des influences subies, la structure morphologique est sujette à changer du tout au tout : le français ou l'anglais ont des structures plus proches de celle du chinois que de celle de l'indo-européen. Il résulte de là que l'examen du type linguistique général ne fournit pas le moyen d'établir un classement généalogique des langues. On en peut tirer tout au plus une indication. C'est ainsi que, malgré la ressemblance de type général, on a cessé de former une famille unissant le turco-mongol au finno-ougrien.

Le système phonique, beaucoup moins stable que le système morphologique, a cependant aussi une certaine fixité et fournit des indications utiles.

En revanche, le vocabulaire est sujet à des innovations brusques et capricieuses. Il traduit des influences de civilisation, plutôt que la continuité linguistique, exprimée par le classement généalogique. Sans doute, il se conserve en général un nombre plus ou moins grand de vieux mots usuels, souvent des verbes tels que ceux signifiant « boire » et « manger », « aller » et « venir », etc. et des adjectifs, souvent aussi des substantifs comme les noms de parenté, d'animaux familiers, etc. Mais ce qui concerne la vie sociale, la vie intellectuelle est sujet à l'emprunt. Les termes de civilisation sont fournis en grande partie par de grandes langues de culture qui servent à de vastes domaines et qui souvent ne sont pas de la même famille que les langues auxquelles elles fournissent des mots.

Ainsi la civilisation grecque, qui a fourni la base de la civilisation occidentale moderne, a donné au latin beaucoup de mots, et surtout le grec a servi de modèle aux lettrés qui ont constitué la langue latine écrite. Directement ou indirectement, le grec a été la source ou le modèle de tous les vocabulaires savants qui se sont établis depuis le début de l'ère chrétienne sur tout le bassin méditerranéen.

L'arabe est la source où s'alimentent tous les vocabulaires des peuples de religion islamique.

L'Inde a exercé sur l'Asie centrale, l'Indo-Chine, la Malaisie une forte action, surtout grâce à l'extension du bouddhisme.

Le persan a, depuis le xe siècle, beaucoup agi en Asie centrale et jusque dans l'Inde.

En Extrême-Orient, c'est le chinois qu'on voit fournir le vocabulaire de civilisation à l'Annam, à la Corée, au Japon, par exemple.

Le nombre de ces grandes langues de civilisation est petit et l'influence de chacune a été immense.

Les alphabets employés indiquent en général quelles sont les influences de ce genre qui se sont exercées. L'extension des

alphabets cunéiformes, grec (et latin, dérivé du grec), araméen, arabe, indien, chinois donne une première indication sur les zones d'influence de chaque langue de civilisation.

En somme les trois principaux types de faits avec lesquels on opère sont en premier lieu la continuité linguistique, qui se manifeste surtout dans le système morphologique et sur laquelle repose toute la classification généalogique — en second lieu, les influences qui ont déterminé les changements (ces influences sont presque toujours non déterminées jusqu'ici, il est vrai, et le présent ouvrage n'en pourra presque pas faire état, mais l'importance en est capitale) — en troisième lieu, l'action des langues de civilisation dont relèvent en grande partie les faits relatifs au vocabulaire.

*
* *

Qu'il s'agisse de l'extension des langues communes, d'où résultent les familles de langues définies, ou des grandes séries d'emprunts, qui caractérisent le type d'influence de culture, les deux faits dominants qu'on vient d'indiquer, à savoir la langue commune initiale et la langue qui fournit les mots empruntés, traduisent des faits de civilisation : la langue qui se répand est, en général, celle du groupe qui apporte le principe d'organisation sociale, et la langue ou les langues qui fournissent les emprunts, sont celles par lesquelles s'exprime la culture.

Ces grands faits sont ceux qu'on observe dans les cas clairs, ceux qui traduisent des influences de civilisation agissant avec force et avec amplitude.

Mais il peut y avoir des situations plus complexes.

Toutes les familles de langues bien établies supposent un même type de développement historique : extension d'une langue commune sur un large domaine, puis différenciation de cette langue, lorsque disparaissent les conditions qui avaient déterminé l'unification. Cette extension suppose l'existence d'une nation possédant une civilisation propre, et ayant conscience de sa force, de son originalité. Elle résulte du prestige qu'a eu cette nation.

Pareil type de faits s'est souvent réalisé dans le passé, et on le voit se réaliser encore aujourd'hui ; car l'extension des langues communes est un trait frappant du monde actuel. Le français commun par exemple remplace les anciens parlers locaux qui s'éliminent. L'anglais s'étend sur le monde. La grande extension des langues indo-européennes tient à ce que, parlées par une nation douée du sens de l'organisation et de la domination, elles se sont peu à peu substituées à un grand nombre d'autres langues, comme on le voit en Italie pour le latin. L'extension des langues sémitiques, de l'assyro-babylonien, de l'hébréo-phénicien, de l'araméen, et, depuis le VII[e] siècle après J.-C., de l'arabe, est un fait qu'on suit historiquement.

Mais il n'est pas évident que les choses se soient partout passées de cette manière. Il peut s'être produit des mélanges de populations plus troubles, en des conditions moins simples. Dès lors, on ne saurait appliquer à toutes les langues les procédés de comparaison qui sont de mise là où il y a eu extension d'une langue commune, comme il est arrivé pour l'indo-européen, le sémitique, le finno-ougrien, l'indonésien, le bantou par exemple. On peut se demander par exemple si les langues américaines, encore mal connues pour la plupart et peu étudiées au point de vue comparatif, se prêteront jamais à l'établissement de grammaires comparées précises et complètes ; les sondages faits jusqu'ici promettent peu.

Dans les langues où le système morphologique ne comporte pas de formations obtenues par des variations compliquées des mots, comme celles d'Extrême-Orient ou certains idiomes soudanais, on peut se représenter de véritables mélanges de langues. Au contraire une langue comme le latin ne peut se mélanger à aucune autre : on emploie ou on n'emploie pas la morphologie latine ; si on l'emploie, on parle latin, quelque vocabulaire qu'on y joigne ; si on ne l'emploie pas, on ne parle pas latin, même quand le vocabulaire est purement latin. Mais il est concevable qu'une langue telle que l'annamite, où toute la grammaire consiste dans la manière de grouper des mots invariables, puisse se mêler à une autre. Par suite, dans les cas de ce genre, non

seulement il est presque impossible d'établir une classification généalogique et de la démontrer, faute d'avoir des données morphologiques par lesquelles seules la preuve se laisse administrer avec rigueur ; mais il n'est même pas évident que le principe de la classification généalogique s'applique. Ce doute doit être signalé, bien que, en fait, partout où l'on a des données suffisantes, les langues – quel que soit le degré de civilisation de ceux qui les parlent — semblent se ramener de plus en plus à un petit nombre d'origines communes, au fur et à mesure qu'elles sont mieux étudiées.

Exprimant des faits historiques de types divers, la classification généalogique des langues ne saurait être homogène. Le sens du mot, parenté de langues, varie donc d'un groupe linguistique à l'autre, d'une langue à l'autre, d'un parler à l'autre. Essayer de faire une classification exacte et complète de toutes les langues en familles rigoureusement définies, c'est montrer déjà qu'on n'a pas compris le principe de la classification généalogique des langues.

En beaucoup de cas, les données dont on dispose sont du reste insuffisantes. Qu'il s'agisse d'Afrique, d'Amérique ou d'Australie, il y a nombre de langues dont la morphologie est mal décrite, ou même ne l'a jamais été. Pour situer ces langues, on est alors réduit à constater des ressemblances de vocabulaire et de type phonique, moins probantes par nature que ne sont les détails grammaticaux définis avec lesquels on démontre la réalité des familles indo-européenne, sémitique, finno-ougrienne, bantou, etc.

Une certaine unité de procédés phoniques indique déjà parfois parenté. Ainsi les langues indo-européennes n'offrent guère les consonnes laryngales, qui sont au contraire courantes en sémitique. Il y a, dans les concordances de ce genre, sinon une preuve solide, du moins une indication. Quand une indication phonique est confirmée par des concordances nettes et systématiques de vocabulaire, portant sur des mots anciens peu sujets à l'emprunt, on peut considérer la parenté comme probable.

Enfin les lacunes qui subsistent dans la classification généalo-

gique des langues tiennent en une large mesure à l'insuffisance des descriptions et des études comparatives faites jusqu'ici ; on n'en pourra mesurer la portée réelle que le jour où les recherches auront été poussées beaucoup plus avant.

En effet, faute d'études complètes et approfondies, bien des parentés de langues restent à établir, qu'on entrevoit seulement aujourd'hui. De plus, beaucoup de langues sont mal décrites ; pour une notable partie des langues africaines, américaines, australiennes, on ne possède que des vocabulaires ; la grammaire en est presque inconnue. Là même où l'on a des descriptions relativement complètes, le travail de comparaison est à peine amorcé le plus souvent.

*
* *

Il n'entre pas dans le plan d'un manuel aussi bref que celui-ci de discuter la question difficile et controversée des rapports entre les grandes familles établies. Il a été produit des faits qui tendent à établir une parenté entre le sémitique et l'indo-européen, entre le finno-ougrien et l'indo-européen. Dans les deux cas, les rapprochements portent bien plus sur le vocabulaire que sur la morphologie proprement dite. L'hypothèse d'une parenté lointaine de l'indo-européen, des anciennes langues asianiques (lycien, carien, etc.), du caucasique, du finno-ougrien et du chamito-sémitique n'a rien d'invraisemblable, et l'on conçoit la possibilité d'une langue commune dont toutes ces langues seraient des transformations diverses. Mais les traits de cette langue commune sont si peu dessinés qu'il serait prématuré de rien affirmer, et, par suite, d'en faire état ici. La démonstration est rendue malaisée par le fait que la morphologie compliquée de l'indo-européen, par exemple, semble résulter d'un développement relativement récent, et que, par suite, on n'a guère de chance d'en retrouver les éléments en sémitique ou en finno-ougrien. Or la parenté de langue perd son sens, on l'a vu, là où ne se reconnaissent pas des continuités de formes grammaticales.

A plus forte raison, il convient de laisser en suspens la ques-

tion de l'unité d'origine du langage humain et, bien entendu, par là même celle de l'origine du langage.

La parité des conditions anatomiques, physiologiques et psychiques dans les divers types humains est telle que les traits essentiels de structure sont sensiblement les mêmes partout : tous les hommes se servent de procédés phoniques semblables en gros, tous parlent par mots groupés de diverses manières. Le détail varie ; le fond des procédés linguistiques est le même dans toute l'humanité. Dès lors, il est malaisé de prouver que quelques ressemblances de détail qui ne se ramènent pas à des systèmes morphologiques définis par des faits particuliers, établissent une parenté entre certaines langues. Ici, on ne fera état que des parentés établies par une grammaire comparée au moins esquissée dans ses grandes lignes ou, à défaut de témoignages précis sur la grammaire, par des groupes cohérents de rapprochements frappants de mots peu empruntables en général et par des ressemblances du type phonétique.

Néanmoins, il apparaît dès maintenant que la plupart des langues se rattachent à un petit nombre de familles définies ; ainsi presque toutes les langues de l'Europe et une partie de celles de l'Asie appartiennent au groupe indo-européen. Toutes les langues du sud de l'Afrique, sauf le hottentot, appartiennent au groupe bantou ; et l'on entrevoit que toutes les langues des populations nègres seraient parentes entre elles. C'est dire que, à des dates historiques ou proches de l'époque historique, un nombre restreint de langues communes s'est étendu. Le fait n'a rien de surprenant : les langues sont faites pour communiquer ; elles rendent d'autant plus de services qu'elles permettent de communiquer avec plus d'hommes. Les hommes ne cessent donc de réagir contre l'émiettement linguistique, qui résulte des événements historiques et de l'emploi des langues.

Du reste, les langues actuellement conservées et dont il subsiste des documents écrits ne représentent pas tous les types qui ont pu exister. On connaît, par des textes plus ou moins étendus ou par quelques témoignages, beaucoup de langues qui ont cessé de se parler au cours de l'époque historique. En Italie, par exemple,

l'étrusque, le vénète, le messapien, le sicule, dont il subsiste des inscriptions, ont disparu, ainsi que les parlers osques et ombriens apparentés au latin. Toutes les anciennes langues de l'Asie antérieure plus ou moins connues par des textes, babylonien, hittite, carien, lydien, lycien, phrygien, ont disparu sans laisser de survivances, et du syriaque il ne subsiste que quelques parlers dispersés. De même qu'il y a des races d'hommes éteintes, bien des familles de langues se sont évanouies, et leur disparition a effacé des transitions qui seraient utiles pour établir les rapports entre les langues conservées.

Les langues qu'on n'arrive pas à grouper peuvent être les débris de groupes dont tous les autres représentants sont sortis de l'usage, et dont rien n'est attesté, hormis une seule langue conservée par hasard.

On sait combien la période historique de l'humanité est courte par rapport aux périodes préhistoriques. La classification généalogique fait ressortir quelques groupements qui résultent d'événements historiques des derniers millénaires de l'humanité. Elles ne permettent même pas d'entrevoir ce qui s'est passé dans des millénaires bien antérieurs, au temps lointain où l'humanité a constitué le langage, instrument essentiel de la vie sociale et du développement industriel et intellectuel.

Si l'on est loin de pouvoir grouper toutes les langues communes, et si surtout la grammaire comparée des groupes, même les mieux connus, est imparfaite, si le groupe indo-européen est le seul où le travail soit très avancé, si hors du sémitique, du finno-ougrien, de l'indonésien et du bantou, presque tout le travail comparatif reste à faire, il va de soi que la question de l'unité d'origine, sinon du langage humain, du moins des langues actuellement connues, ne peut être abordée d'une manière utile.

*
* *

La terminologie linguistique est peu fixée et peu précise : il y a lieu de définir les termes de *langue* (idiome), *dialecte*, *parler*, *patois*, *langue spéciale*, *argot*.

Le mot *langue*, au sens large, est le plus général : il désigne tout un ensemble de moyens linguistiques employés par un groupe d'hommes, quelle que soit l'étendue de ce groupe et quelle qu'en soit la valeur au point de vue de la civilisation. Il faut compter autant de *langues* (on dit aussi *idiomes*) qu'il y a d'ensembles de ce genre, assez différents pour que les divers sujets parlants qui les emploient ne se comprennent pas sans apprentissage.

Cette limite d'intelligibilité est absolue en beaucoup de cas ; un sujet parlant chinois, arabe ou anglais est inintelligible pour un sujet connaissant seulement le français. Mais un sujet employant la manière de parler en usage dans un village picard est à demi intelligible pour un Parisien. Entre l'inintelligibilité absolue et l'intelligibilité absolue, il y a tous les degrés possibles. Dans un groupe étendu, tel qu'est celui du groupe employant des parlers locaux du type français, les gens de localités voisines s'entendent d'une manière plus ou moins complète, alors que des sujets de régions éloignées ne s'entendent nullement ; ainsi un sujet parlant normand n'entend pas un sujet employant un parler local franc-comtois ou wallon.

On entend par *parler* l'ensemble des moyens linguistiques employés par un groupe local à l'intérieur d'un groupe occupant une aire étendue. Il y a les parlers français dans la France septentrionale, les parlers provençaux, gascons, etc. dans la France méridionale, les parlers italiens en Italie, etc. Les parlers locaux sont souvent nommés *patois*.

Dans la mesure où le mot *langue* s'oppose au mot *parler*, il désigne un ensemble de procédés définis, et dont on a conscience, en somme. Ainsi, chacun des parlers français est un ensemble d'usages locaux, qui varient sensiblement suivant les individus ; ces usages oscillent autour d'une moyenne, et ils offrent en gros une régularité, suivant des règles d'usage qui sont souvent très fixes ; mais ces règles ne sont codifiées nulle part. Au contraire, le français est une langue commune qui a des règles rigoureusement fixées, une tradition littéraire. Le français s'oppose ainsi aux parlers français. Ainsi entendu, le mot *langue* a un sens très élastique.

La pratique comporte un flottement suivant les classes sociales, suivant les conditions d'emploi (langue parlée et langue écrite), etc.

A l'intérieur d'un groupe linguistique étendu, on constate, en général, que certains parlers offrent des traits communs et que les sujets parlants de certaines régions, ont le sentiment d'appartenir à un même sous-groupe : en pareil cas, on dit que ces parlers font partie d'un même *dialecte*. On peut ainsi grouper les parlers français : parlers français proprement dits (Ile de France), picards, wallons, normands, champenois, lorrains, bourguignons, franc-comtois, etc. Cette notion est fuyante. On ne peut le plus souvent assigner de limite précise au dialecte, parce que chacune des particularités, par lesquelles se caractérisent les dialectes, a en général une limite propre, différente de celle de toute autre particularité. Un dialecte se reconnaît à ce que les limites d'un certain nombre de particularités, sans concorder exactement, sont proches les unes des autres, si bien que ce groupe de parlers constituant un dialecte offre des traits communs.

Parmi les sujets employant une même langue ou un même parler, il existe souvent — en dehors des différences locales — des groupements d'individus qui offrent certaines particularités linguistiques : les gens qui exercent une même profession, qui font partie de mêmes groupements peuvent constituer des *langues spéciales*. D'ordinaire ces langues ne se distinguent pas de la langue générale par la prononciation ou par la morphologie. Les particularités portent sur le vocabulaire presque uniquement.

Les vocabulaires de petits groupes, doublant plus ou moins largement le vocabulaire de la langue commune, sont souvent nommés *argots* ; ce terme s'emploie surtout pour les parlers de malfaiteurs ou de gens suspects ; mais on en étend souvent aussi le sens ; et l'on parle de l'*argot* de la caserne, de telle ou telle école, etc.

En outre, il y a *langue spéciale* là où une profession, un groupement particulier exigent l'emploi de termes techniques ; c'est ce qui arrive dans les divers métiers, dans les sports, etc.

Enfin, il faudrait tenir compte des formes altérées que

prennent des langues employées d'une manière imparfaite par des gens qui ne cherchent pas à les parler d'une manière normale. Ainsi les esclaves nègres des colonies n'ont pas cherché à parler normalement le français ou l'espagnol de leurs maîtres : les parlers « créoles » usités — espagnols ou français — sont ainsi de l'espagnol ou du français, privés de presque toute leur grammaire, affaiblis dans leur prononciation, réduits à un petit vocabulaire. En Extrême-Orient, des gens de parlers divers se servent, notamment dans leurs relations avec les Européens, d'un anglais dégénéré qu'on appelle le *pidgin-english*. Dans le bassin de la Méditerranée, il y a des langues, ainsi réduites, qu'on nomme des *sabirs*.

Les langues de cette sorte ont un minimum de grammaire, et elles sont sujettes à toutes sortes de mélanges. Si elles venaient à se fixer dans l'usage, elles auraient une place à part dans la classification généalogique : fondée sur l'histoire, cette classification doit présenter les rapports de manières diverses, là où les rapports sont divers.

Il est naturellement impossible de faire entrer dans la classification les langues artificielles, telles que le volapük, l'espéranto, l'ido. Ces langues — dont aucune n'est vraiment entrée dans l'usage — n'ont du reste obtenu de succès que dans la mesure où elles sont une moyenne entre des langues parlées par ceux à qui elles sont destinées. Le vocabulaire de celles des langues qui semblent susceptibles de servir de base à un développement réel, l'espéranto et l'ido, est fait systématiquement avec des termes communs au plus grand nombre possible des grandes langues européennes ; il en résulte que, en fait, leur vocabulaire est surtout latin et que ces langues artificielles sont proches des langues romanes. Ainsi on revient toujours à la classification généalogique.

*
* *

On verra, par les exposés qui suivent, combien la classification est encore imparfaite, et pourquoi elle est destinée à le

demeurer. Pour faire une bonne classification généalogique, il faudrait pouvoir suivre l'histoire linguistique de chaque domaine, en marquant toutes les influences subies, depuis le moment où le langage s'est constitué. Or, même dans les cas les meilleurs, on entrevoit à peine quelques moments récents de cette histoire. Ainsi, pour la France, on sait que le gaulois s'est propagé par l'effet de la conquête celtique, qui date du premier millénaire avant l'ère chrétienne, que le latin s'y est introduit durant les premiers siècles de l'ère chrétienne. Mais on ne sait de quels idiomes le gaulois a pris la place, ni quelles langues autres que le gaulois subsistaient lors de la conquête romaine. Quant aux langues parlées avant la conquête gauloise, on en ignore tout, et le fait que beaucoup de noms de lieu de la Gaule ne se laissent pas interpréter par le celtique et ne sont sans doute pas celtiques, est le seul témoignage linguistique attestant l'existence en Gaule de parlers antérieurs au gaulois.

La grammaire comparée des langues indo-européennes fournit à l'ensemble de la linguistique historique un modèle à imiter. En dépit de lacunes immenses et d'insuffisances visibles, ce qui a été obtenu sur ce domaine demeurera sans doute un idéal inaccessible dans la plupart des autres cas. Mais il reste beaucoup à trouver partout. Dans le présent livre il y a donc un programme de recherches plus qu'une somme de résultats. Le travail qui reste à faire est immense, et il ne pourra être fait que si des ressources considérables sont mises aux mains des linguistes.

A. MEILLET.

BIBLIOGRAPHIE

Les deux thèses opposées dans la théorie de la classification généalogique sont exposées l'une, du point de vue du développement des langues communes dont le sujet parlant a conscience, dans :

A. MEILLET, *Linguistique historique et linguistique générale*,, Paris 1921, p. 76-109 (en tenant compte des chapitres précédents et suivants) ;

l'autre, en partie du point de vue des langues « mixtes », dans :

H. SCHUCHARDT, *Sprachverwandtschaft*, paru en 1917 dans les *Sitzungsberichte* de l'Académie de Berlin ; reproduit maintenant, pour l'essentiel, dans

Hugo Schuchardt-Brevier, Halle, 1922, p. 167 et suiv. : on tiendra compte de ce qui précède et de ce qui suit dans le recueil Schuchardt.

Ces idées sont discutées avec sagacité dans :

J. Vendryes, *Le langage*, Paris 1921, p. 349 et suiv.

Les classifications a priori sont indiquées et critiquées brièvement, mais de manière suffisante, dans un ouvrage fait malheureusement pour soutenir la thèse de l'unité d'origine des langues :

Trombetti, *Elementi di glottologia*. Bologne, 1922-1923, p. 6 et suiv.,

et dans :

O. Jespersen, *Language*, Londres, 1922, p. 76 et suiv.

Le livre de M. Jespersen comporte une discussion, très personnelle, de presque tous les faits envisagés dans cette introduction.

On devra aussi tenir grand compte de :

E. Sapir, *Language*, New York, 1921, p. 204-235.

LANGUES INDO-EUROPÉENNES

Généralités.

La famille indo-européenne [1] est celle à laquelle étaient réservées dans l'histoire les destinées les plus hautes. Elle a créé les formes linguistiques qui ont servi aux littératures les plus belles et les plus riches ; elle a aidé à se répandre les civilisations qui ont conquis le monde. Il se parle aujourd'hui des langues indo-européennes dans la presque totalité de l'Europe et de l'Amérique, dans une grande partie de l'Asie, dans une partie notable de l'Afrique et de l'Océanie.

Ce succès prodigieux, si on le compare au sort de tant d'autres langues, qui ont végété sans éclat dans des territoires resserrés ou même ont été étouffées par des langues plus vivaces, s'explique par des causes historiques au nombre desquelles il faut compter le génie propre des Indo-européens, le caractère de leurs institutions, le prestige de leur civilisation. Partout où ils se sont étendus, les Indo-européens représentent une aristocratie conquérante, dominant des peuples vaincus et leur imposant avec son autorité l'usage de sa langue. L'extension des langues indo-européennes s'est faite par une série de conquêtes. Le sanskrit a pénétré en conquérant dans l'Inde, comme le latin en Italie,

1. Consulter avant tout K. Brugmann, *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*, Strassburg, 2e édition, 1897-1916, et A. Meillet, *Introduction à l'étude comparative des langues indo-européennes*, Paris, 5e édition (1922). Un abrégé du *Grundriss* de Brugmann, publié à Leipzig en 1902-1904, a été traduit en français sous la direction de A. Meillet et R. Gauthiot (Paris, Klincksieck, 1905). Bibliographie des dernières années, dans *Indogermanisches Jahrbuch*, Strassburg, puis Berlin, 1914 et suiv.

comme l'allemand dans la partie orientale de l'Allemagne actuelle, comme le russe en Russie, comme le français, l'espagnol, l'anglais dans le nouveau monde. Un des résultats de la conquête était d'étendre la langue des conquérants sur un territoire plus ou moins vaste. Cette langue y restait sensiblement uniforme jusqu'au jour où le relâchement des liens communs (politiques, économiques, religieux ou autres) permettait la constitution de groupes distincts et laissait à chacun de ces groupes la liberté d'un développement indépendant. Ainsi la ruine de l'empire romain a détruit l'unité de la langue latine ; réduit à l'état de langue morte, le latin céda la place aux langues néo-latines, qui se partagèrent le vaste domaine sur lequel régnait uniformément le latin.

L'unité linguistique à laquelle on donne le nom d'indo-européen représente une unité de civilisation, résultant de la conquête, comme c'est le cas du latin ; et les langues indo-européennes, comme les langues néo-latines, doivent leur naissance à des événements, d'ailleurs inconnus, à la suite desquels les liens qui maintenaient l'unité se seront brisés. Cette conclusion n'exclut pas l'hypothèse qu'il y ait eu sur le domaine indo-européen des divisions dialectales [1] ; il est vraisemblable au contraire que l'unité de la langue commune enfermait des éléments de différenciation, auxquels la rupture de l'unité ne fit que donner un libre essor. On doit admettre d'autre part que la dislocation de l'unité indo-européenne ne s'est pas faite en un seul jour ni de la même façon sur tous les points du domaine. Cela explique en partie les différences de structure et les variétés de tendances que l'on observe parmi les langues indo-européennes dès la date la plus ancienne où elles sont attestées.

L'unité indo-européenne a certainement été précédée d'un développement linguistique, qui l'a préparée. De ce fait se dégagent deux ordres de conséquences. C'est d'abord qu'il peut y avoir des groupes linguistiques apparentés à l'indo-européen et qui se distingueraient de lui comme le français de l'italien, le

1. A. Meillet, *Les dialectes indo-européens*, Paris, 1908 réimpr. 1922.

latin du gaulois, l'italo-celtique du grec. On a en fait essayé en ces derniers temps de rattacher à l'indo-européen tantôt le finno-ougrien et tantôt le sémitique [1]. L'hypothèse que ces trois familles de langues seraient issues d'un ancêtre commun n'est pas dénuée de vraisemblance ; il faut seulement laisser à de nouvelles recherches le temps de la démontrer.

La seconde conséquence est qu'il a pu y avoir des langues indo-européennes qui aient pris une extension propre antérieurement à la constitution de l'unité indo-européenne et dont le développement se soit fait indépendamment de cette unité. Ces langues seraient à l'indo-européen ce que l'osque et l'ombrien sont au latin de Rome de l'époque républicaine, l'achéen au dorien du Péloponnèse ou mieux encore les vieux dialectes grecs à la *koiné* hellénistique. Ainsi peut-être se place le pseudo-hittite par rapport à l'indo-européen [2]. Dans la mesure où l'on peut se fier à l'interprétation proposée récemment pour les textes de ce pseudo-hittite, il apparaît qu'environ quinze cents ans av. J.-C., dans la région orientale de l'Asie Mineure, était parlée une langue d'un type voisin de l'indo-européen, mais ne se rattachant à aucun des groupes indo-européens connus et présentant une structure trop aberrante pour être ramenée à un stade d'évolution commun à chacun de ces groupes. L'avenir réserve peut-être aux chercheurs la découverte d'autres langues comparables au pseudo-hittite.

1. Sur les rapports de l'indo-européen et du finno-ougrien, voir une bibliographie dans l'ouvrage de S. Feist, *Kultur, Ausbreitung und Herkunft der Indogermanen*, p. 396 note ; mais consulter surtout Hermann Jacobsohn, *Arier und Ugrofinnen*, Göttingen, 1922. Sur ceux de l'indo-européen et du sémitique, consulter Hermann Möller, *Semitisch und Indogermanisch*, Copenhague, 1906, et *Indo-europaeisk-semitisk sammenlignende glossarium*, Copenhague, 1909 ; Pedersen, *Indog. Fschg.*, XXII, 341-365, et A. Cuny, *Revue des études anciennes*, t. XIV à XXII.

2. Marstrander, *Caractère indo-européen de la langue hittite*, Christiania, 1919 ; F. Sommer, *Hethitisches*, Leipzig, 1920 ; A. Debrunner, *Die Sprache der Hethiter*, Berne, 1921 ; les mémoires de Joh. Friedrich et E. Forrer, *Zeitschrift der deutschen morgenländischen Gesellschaft*, Neue Folge, Band I, Heft 2, p. 153-269 ; Herbig, dans *Indogermanisches Jahrbuch*, VIII (1922), p. 1-20.

Même si les deux conséquences du principe posé ne pouvaient pas recevoir de vérification d'un examen plus attentif des faits, elles devraient toujours s'imposer à l'esprit du linguiste à titre d'hypothèses. Elles offrent le double avantage de placer l'indo-européen dans les conditions historiques où se sont trouvées au cours des âges les grandes langues communes qui se sont constituées et d'écarter l'idée que l'indo-européen serait apparu un beau jour comme un produit spontané sans aucun lien avec les langues des communautés humaines contemporaines ou antérieures.

Après ce qui vient d'être dit, il peut paraître oiseux de rechercher où était l'habitat des Indo-européens [1]. On a prétendu le fixer tantôt en Asie centrale, dans le voisinage du Turkestan actuel, tantôt en Europe orientale, soit dans les plaines de la Russie méridionale, soit plus au Nord en direction de la Baltique. La question, mainte fois discutée, a pris en ces derniers temps un regain d'actualité. Ethnographes et préhistoriens, archéologues et linguistes mettent en commun pour la résoudre les ressources de leurs sciences respectives. Le résultat laisse toujours à désirer. On peut tirer de la linguistique quelques précisions sur l'organisation sociale, la religion, le genre de vie, l'outillage des Indo-européens. Sur la question de l'habitat, ce que fournit la linguistique est maigre : l'étude comparative des vocabulaires permet de conclure que l'indo-européen commun se parlait dans une région froide, éloignée de la mer, où poussaient le bouleau, le chêne et le hêtre, où vivaient l'ours et le loup. Ce sont là des données bien vagues. Elles n'intéressent guère le linguiste. Peu importe la région où la nation indo-européenne avait établi une unité de langue. Cette région d'ailleurs a pu varier, suivant les

1. Ch. de Ujfalvy, *Le berceau des Aryas d'après les ouvrages récents*, 1884 ; S. Reinach, *L'origine des Aryens*, 1892 ; G. Sergi, *Gli Arii in Europe e in Asia*, 1902 ; E. de Michelis, *L'origine degli Indo-Europei*, 1903 ; H. Hirt, *Die Indogermanen, ihre Verbreitung, ihre Urheimat und ihre Kultur*, Strassburg, 1905-1907 ; M. Much, *Die Heimat der Indogermanen im Lichte der urgeschichtlichen Forschung*, Iena, 2e éd. 1904 ; O. Schrader, *Die Indogermanen*, Leipzig, 1911 ; S. Feist, *Kultur, Ausbreitung und Herkunft der Indogermanen*, Berlin, 1913.

déplacements des peuples, au cours même de la période d'unité. L'essentiel est d'atteindre par la méthode comparative l'unité linguistique d'où les langues indo-européennes sont issues.

Le grand intérêt qu'offre le développement de ces langues est dans la variété des transformations linguistiques qu'il présente. Bien que les plus anciens documents que l'on en possède soient d'une date relativement basse si on les compare aux plus anciens textes du babylonien ou de l'égyptien, ils offrent une structure extrêmement archaïque par rapport aux formes ultérieures des mêmes langues. Entre l'indo-européen, tel que la comparaison permet d'en reconstituer les traits essentiels, et la plupart des langues indo-européennes parlées aujourd'hui, il y a de très grandes différences de structure.

Des différences non moins grandes apparaissent entre les diverses langues issues de l'indo-européen commun. Ces différences peuvent tenir sans doute au développement intrinsèque de ces langues qui ont chacune fait prévaloir des tendances propres : on comprend sans peine que le vieux slave se distingue fortement du latin ou le grec du germanique commun par le seul fait que ces langues ont poursuivi pendant plusieurs siècles des développements indépendants. Mais il faut reconnaître l'influence la plus décisive aux actions extérieures que les groupes différenciés ont subies chacun de leur côté et qui en ont profondément modifié le caractère initial. Comme ces actions ne se sont produites ni à la même date, ni avec la même ampleur et qu'elles ont atteint les langues de façon inégale, on observe des différences frappantes entre le mode et le degré de transformation de deux dialectes originairement voisins lorsqu'on les considère à un même moment de l'histoire peu de siècles après leur séparation.

C'est ainsi que l'indien et l'iranien, le brittonique et le gaélique, dont la séparation respective est bien postérieure à la dislocation de l'unité indo-européenne, ont entre eux, lorsqu'ils apparaissent dans l'histoire, des différences notables. Dans chacun des membres de chaque groupe, la morphologie, le vocabulaire ont été altérés différemment, et la transformation a été plus ou

moins rapide. Dès les plus anciens textes, le brittonique offre à beaucoup d'égards un état linguistique plus évolué que n'est celui de l'irlandais moderne. Entre toutes les langues indo-européennes, l'iranien a eu l'évolution la plus précoce ; les parlers iraniens se trouvent au début de l'ère chrétienne à un niveau linguistique sensiblement égal à celui que les langues romanes devaient atteindre dix siècles plus tard. C'est qu'en se répandant sur un vaste domaine compris entre le golfe Persique et la Caspienne, le Pamir et la région au nord de la Mer Noire, l'iranien rencontrait devant lui de vieilles langues de civilisation ; il a subi plus tôt que n'importe quel autre groupe indo-européen les actions extérieures les plus fortes et les plus variées.

Inversement une langue comme le lituanien offre certaines particularités remarquablement archaïques ; des traits y sont conservés jusqu'à nos jours, que des langues comme le grec ancien ou le latin avaient laissé perdre. C'est que le lituanien, langue rurale, à l'abri des influences extérieures et à l'écart des grands courants de civilisation, est parlé depuis de longs siècles par une population stable, sur un domaine étroitement limité. Quelle différence avec une grande langue de civilisation comme le français ou l'anglais ! Il n'y a pas commune mesure entre le lituanien et l'anglais, parce que la valeur de chacune de ces langues est différente : et il est fatal que la différence de valeur entraîne des différences de structure.

Malgré la variété des péripéties et des résultats du développement de chaque langue indo-européenne, il y a cependant dans ce développement certaines tendances communes qui dominent. Ces tendances devaient aboutir dans toutes les langues à des transformations analogues. On peut en effet, lorsqu'on les a dégagées, distinguer en général deux types de langues indo-européennes : un type ancien représenté surtout par le sanskrit, l'iranien ancien, le grec ancien et le latin ; un type moderne représenté par les langues postérieures à l'ère chrétienne et notamment par les langues romanes, germaniques et celtiques, par les parlers iraniens depuis le IIIe s. ap. J.-C. et les parlers arméniens. Les principales différences des deux types apparaissent si l'on passe en

revue les traits généraux de la structure de l'indo-européen et les principes qui tendaient à la transformer.

La phonétique de l'indo-européen [1] était caractérisée par l'absence de phonèmes spirants ou chuintants ; seule parmi les consonnes, la sifflante *s* représentait les continues. En revanche, il y avait un jeu complexe d'occlusives, sourdes et sonores, qui pouvaient être aspirées ou non et qui se répartissaient en labiales, dentales, prépalatales et vélaires. Le vocalisme était simple : en dehors de *e* et de *o* (brefs ou longs, sans distinction de timbre), il ne comportait guère qu'un *a* (bref ou long). Le trait le plus original du phonétisme indo-européen était l'existence d'une troisième catégorie de phonèmes, les sonantes (semi-voyelles, liquides ou nasales), qui pouvaient jouer suivant le cas le rôle de voyelles, de consonnes ou de seconds éléments de diphtongue. Or, on observe partout des tendances générales à la simplification du système des occlusives, à la création de spirantes (par affaiblissement de l'articulation des occlusives), à l'élimination du système des sonantes (qui deviennent définitivement voyelles ou consonnes), à la réduction des diphtongues. Le système vocalique s'enrichit en outre par des variations de timbre, liées généralement à des différences de quantité, et par le développement de la voyelle *a*, issue d'origines variées. L'accident le plus grave a été causé par la transformation du rythme de la langue. Le rythme indo-européen était purement quantitatif et l'accent, qui était de hauteur, n'exerçait aucune action sur la quantité des voyelles. Mais peu à peu, à des dates variables suivant les langues, le rythme quantitatif s'altère, l'accent tend à prendre une place fixe dans le mot et, sous forme d'accent d'intensité, à régler les rapports quantitatifs. La structure phonétique des mots a été de ce fait plus ou moins complètement transformée ; les syllabes finales notamment ont été souvent gravement atteintes.

La morphologie indo-européenne [2] reposait sur le principe que

1. Voir A. Meillet, *Introduction*, p. 57.
2. Voir A. Meillet, *Introduction*, p. 115.

tout mot portait en lui-même la marque de son emploi : par suite, la phrase était constituée d'éléments autonomes et indépendants. Il y avait bien aussi des mots accessoires (particules, adverbes, etc.), même en nombre important ; mais leur rôle était seulement de marquer la liaison des phrases ou bien d'ajouter des précisions (de lieu, de temps, de circonstances) au sens qu'exprimaient les autres mots. Ceux-ci, sans le secours des mots accessoires, avaient toujours un sens complet et un emploi défini.

Tout mot indo-européen comprenait trois éléments — racine, suffixe, désinence — soumis chacun à un jeu très subtil d'alternances vocaliques (*e*, *o*, zéro) pourvues de signification, le tout complété par le ton dont la place avait également une valeur significative. L'ensemble de ces caractères complexes conférait au mot une plénitude absolue ; car aucun des éléments qui le constituaient n'avait d'existence indépendante et ne pouvait être employé isolément. L'élément essentiel était la racine, qui exprimait le sens du mot, tandis que les suffixes et les désinences, aussi bien que le degré vocalique et la place du ton, indiquaient les catégories morphologiques et les modalités de l'emploi. La racine n'avait qu'une existence idéale, dans l'esprit de ceux qui parlaient ; mais c'est d'elle que les mots étaient formés directement, indépendamment les uns des autres. La grammaire ne comportait donc pas des systèmes de paradigmes ordonnés, à la façon d'une grammaire latine ou d'une grammaire russe. Cependant, bien qu'il n'y eût pas un système cohérent du verbe ou du nom, il y avait entre les formations verbales et les formations nominales des différences fondamentales ; les unes et les autres avaient des éléments suffixaux, des désinences qui leur étaient propres, et suivaient même souvent des règles syntaxiques particulières. Ces différences tenaient sans doute à l'opposition de deux types de phrases (phrase verbale et phrase nominale), que l'indo-européen distinguait rigoureusement.

Dès le début du développement des langues indo-européennes, on constate une tendance à diminuer le rôle de la racine, en y substituant une forme immuable, le radical, pris comme élément fondamental du mot. Le jeu des alternances se limite et se sim-

plifie. En même temps le système du verbe et le système du nom tendent à s'organiser. C'est pour le verbe que l'organisation a été la plus rapide et la plus solide : le latin et le germanique, le baltique et le slave se créent de bonne heure un verbe à deux thèmes, dont les éléments sont si fermement ordonnés qu'ils subsistent sans changement appréciable jusque dans les langues parlées aujourd'hui. Le système du nom n'a jamais été aussi ferme : la tendance générale est à la simplification de la flexion. Les cas concrets perdent de leur importance par suite de l'emploi de particules et notamment d'anciens éléments adverbiaux transformés en prépositions. La création d'articles, qui s'est opérée indépendamment dans bon nombre de langues, sert à l'expression de rapports intéressant la valeur et l'aspect du nom. L'emploi du nom comporte ainsi, beaucoup plus tôt que celui du verbe, l'addition d'éléments accessoires indispensables.

La transformation du système morphologique ancien a été précipitée par un accident phonétique, l'altération des finales. Des finales affaiblies, parfois même complètement amuies, privaient le mot des marques de son emploi. Les langues ont remédié à cette situation par divers procédés, dont le plus simple était la création de nouvelles désinences, mais dont le plus répandu a été le développement des outils grammaticaux. Ceux-ci augmentent en nombre et en importance à mesure que la flexion se simplifie davantage. Là où la flexion est réduite au minimum, comme en anglais ou en persan, les rapports grammaticaux ne peuvent être exprimés que par des moyens accessoires, extérieurs au mot (emploi de particules, de mots auxiliaires, attribution d'une valeur grammaticale à l'ordre des mots). Le mot lui-même tend à devenir alors un symbole abstrait, également apte à l'expression d'une notion nominale ou verbale, et ne portant en lui-même aucune marque d'un emploi quelconque. Rien n'est plus éloigné de ce que représentait un mot indo-européen.

Si l'on veut saisir dans leur ensemble l'importance des transformations qu'a subies le type primitif, il suffit de comparer une phrase de grec ancien ou de latin à une phrase de français ou d'anglais modernes. Quant aux détails de la structure, on les

trouvera étudiés dans les divers chapitres de l'*Introduction* de M. Meillet.

*
* *

En Asie, la famille indo-européenne est avant tout représentée par deux grands groupes, l'indien et l'iranien, connus tous deux longtemps avant l'ère chrétienne. Ils ont des rapports linguistiques si étroits qu'on les réunit souvent sous le nom d'indo-iranien ou d'*aryen*. Le mot *arya-* — dont le nom de l'Iran est une forme récente — est en effet celui par lequel les ancêtres communs des Indiens et des Iraniens se désignaient eux-mêmes.

Si le mot *aryen* a été souvent employé pour désigner l'ensemble de l'indo-européen, c'est par un abus de terme auquel les linguistes ont entièrement renoncé.

GROUPE INDO-ARYEN.

Les plus anciens documents du groupe « aryen » de l'Inde ou indo-aryen sont les textes védiques. Ils sont écrits en SANSKRIT[1]. La date en est impossible à fixer; par le fond comme par la forme, ils se révèlent comme d'une haute antiquité. Avant d'être écrits, ils ont dû se transmettre par tradition orale; aussi le *sanskrit védique*, bien que fixé scrupuleusement comme langue religieuse, n'est pas une langue pure. Le fonds principal du sanskrit védique est le dialecte de l'extrémité nord-ouest de l'Inde, du Pendjab; mais il y a été mêlé des éléments provenant de parlers situés à l'Est du domaine. Le plus archaïque des textes védiques est le *Ṛig-Veda* (« Veda des chants »), recueil liturgique d'hymnes; l'*Atharva-Veda* (« Veda des Atharvans, sorte de prêtres »), qui est surtout un recueil de formules de sorcellerie, prières, conjurations et incantations, a des formes sensiblement plus récentes. Bien plus jeunes sont les textes en prose, commentaires sur le Veda (*Brāhmaṇa*) ou traités philosophiques (*Upaniṣad*), ainsi

1. J. Wackernagel, *Altindische Grammatik*, t. I, Einleitung (Göttingen, 1896).

que les grandes épopées en vers (*Mahâbhārata* et *Rāmāyaṇa*), dont la langue est déjà du sanskrit classique.

L'emploi du sanskrit comme langue littéraire suppose qu'il a perdu son caractère de langue religieuse spéciale. Cette transformation ne s'est produite qu'assez tard ; aussi est-il probable que les grandes épopées sanskrites existaient déjà, sous une forme d'ailleurs inconnue, avant d'avoir été transcrites en sanskrit. La plus ancienne inscription sanskrite, trouvée à Girnar, date seulement de 150 av. J.-C. ; elle est l'œuvre d'un étranger, le roi des Çakas Rudradāman. C'est à partir du IVe siècle de notre ère que le sanskrit devient la langue unique de l'épigraphie officielle. Fixé comme langue savante par le travail de nombreux grammairiens, dont le plus célèbre est Pāṇini (IVe s. av. notre ère), le *sanskrit classique* s'est maintenu jusqu'à l'époque moderne à l'état de langue littéraire et scientifique : une littérature des plus vastes (lyrique, narrative, dramatique, philosophique, technique, etc.) est écrite en sanskrit. Aujourd'hui encore le sanskrit peut à l'occasion servir à des paṇḍits des différentes parties de l'Inde à communiquer entre eux, à peu près comme on employait le latin dans l'Europe du moyen âge. Malgré la fixation rigide de la prononciation et des formes grammaticales, le sanskrit n'a jamais été complètement soustrait à l'action des parlers vivants. Ceux-ci ont constamment fourni des éléments nouveaux au vocabulaire, à mesure que l'usage du sanskrit comme langue profane s'étendait vers l'Est et le Sud.

Bien avant que le sanskrit cessât d'être exclusivement réservé aux usages religieux, il s'était développé dans l'Inde des langues communes, d'extension plus ou moins considérable et qui dérivaient toutes d'une forme linguistique peu différente de celle dont le sanskrit védique est sorti. Ces langues communes sont ce qu'on appelle des PRÂKRITS [1]. Mais les premiers documents datés des langues aryennes de l'Inde sont déjà du moyen-indien. Ce sont les inscriptions du roi Asoka, qui vivait au IIIe s. av. J.-C. On les trouve disséminées dans les régions les plus diverses de

1. R. Pischel, *Grammatik der Prakrit-Sprachen* dans le *Grundriss der indo-arischen Philologie*, Strassburg, 1900.

l'Inde, offrant suivant les lieux de grandes différences dialectales ; elles présentent un état de langue fort évolué par comparaison avec le sanskrit. Les prâkrits sont en général désignés par un nom de région. On distingue ainsi la *çaurasenî*, langue du Çūrasena (région de Mattra), la *mâgadhî*, langue du Magadha (région de Patna), la *mâhârâṣṭrî* langue du Mahârâṣṭra (pays des Mahrattes), etc. ; cela indique en gros la région d'où ils étaient originaires. Mais il n'y a pas nécessairement pour cela de dialecte local à la base de chaque prâkrit. Les prâkrits sont des langues littéraires, amendées et fixées artificiellement par des théoriciens, quelque chose comme le dorien des chœurs de la tragédie attique. Leurs particularités ne sont ni complètement irréelles, ni complètement conformes à la réalité du langage parlé [1]. Certains offrent un mélange réglé par convention et où les souvenirs du sanskrit dominent. Ainsi la *paiçâcî*, bien que fondée, semble-t-il, sur une langue du Nord-Ouest, n'en reproduit que quelques traits dialectaux ; on sent que Guṇâḍhya, qui passe pour l'avoir fixée, ne perdait pas de vue la norme du sanskrit.

Indépendamment des prâkrits, les parlers locaux (*deçabhâṣâ* ou *grāmyabhâṣâ* « langue régionale, langue locale ») continuaient à se développer. Ils ont parfois donné naissance à ce type de langue écrite que les grammairiens de l'Inde nomment *apabhraṃça* [2]. Bien qu'il soit parfois confondu avec les prâkrits, l'*apabhraṃça* paraît avoir été à l'origine une forme intermédiaire entre le parler local et les prâkrits, une sorte de dialecte épuré et normalisé par suite d'une extension à plusieurs patois. Il a pu aussi comme tel servir de langue littéraire.

A côté des langues littéraires, il a existé en moyen-indien des langues religieuses. La plus célèbre est le PÂLI [3]. C'est en pâli que

1. Sur les prâkrits et notamment sur la paiçâcî, voir F. Lacôte, *Essai sur Guṇāḍhya* (Paris, 1908), p. 58.

2. R. Pischel, *Materialien zur Kenntniss des Apabhraṃça*, Berlin, 1902 (fait partie des *Abhandlungen* de la Société des sciences de Göttingen). H. Jacobi, *Bhavisatta Kaha*, München, 1918 ; *Sanatkumāra-Caritam*, München, 1921 (tous deux dans les *Abhandlungen* de l'Académie Bavaroise des Sciences).

3. Burnouf et Lassen, *Essai sur le pâli* ; V. Henry, *Précis de grammaire pâlie*, 1904.

fut rédigé antérieurement au Ier s. av. notre ère le canon bouddhique conservé à Ceylan. Le pâli est certainement d'origine continentale, mais il est difficile d'en localiser la source, car il n'est pas homogène. Ses éléments essentiels paraissent provenir de la région du Mâlva, au Nord d'Indore, mais beaucoup d'éléments hétérogènes s'y sont ajoutés. Il s'est conservé comme une des langues religieuses du bouddhisme. De son côté, le jaïnisme s'est constitué aussi des langues religieuses; la principale, née sans doute dans le Magadha, où fut rédigé le canon jaïnique, fut transportée plus tard, lorsque le centre de cette religion se déplaça, dans le Dekkan et au Guzrate ; elle prit alors certains caractères qui la rapprochent de la *mâhârâṣṭrî*.

On parle aujourd'hui indo-aryen [1] dans l'Inde sur un domaine continu, limité au Nord-Ouest par le domaine iranien, au Nord et au Nord-Est par le domaine tibéto-birman, à l'Est par le domaine munda, au Sud par le domaine dravidien ; la limite de ce dernier s'arrête à peu près au Concan et au bassin moyen de la Godavari. Il faut joindre au domaine aryen de l'Inde la moitié méridionale de l'île de Ceylan.

Sur ce vaste domaine, où l'indo-aryen est parlé par environ 231 millions d'individus, on peut distinguer plusieurs groupes dialectaux. A l'extrémité nord-ouest dans la région montagneuse qui touche au Pamir, se parlent des dialectes assez aberrants et qui d'ailleurs divergent fort entre eux ; on les appelle parfois dialectes himalayens. Ce sont surtout : à l'Ouest, le *véron* et les dialectes du Kafiristan ; au centre, le *khowar*, parlé dans la vallée du Chitral ; à l'Est, les dialectes du haut cours de l'Indus (par exemple le *china*), du Kohistan et du Cachmire. Le *cachmirien* (*kaçmiri*), qui est le mieux connu de ces dialectes, a servi de langue littéraire ; il est pénétré d'éléments iraniens et sanskrits.

1. J. Beames, *A comparative grammar of the modern Aryan languages of India*, 3 vol., London, 1872-1879 ; G. A. Grierson, *The Languages of India*, Calcutta, 1903 ; id , *The Linguistic Survey of India and the Census of 1911*, Calcutta, 1919 ; id., *Indo-Aryan Vernaculars*, dans le *Bulletin of the School of oriental Studies*, London, I, II (1918), p. 47-81, et I, III (1920), p. 51-85.

Les dialectes de l'Inde proprement dite se ramènent aux principaux groupes suivants (voir la planche 9).

Groupe occidental. En descendant le cours de l'Indus, on rencontre d'abord le *lahnda* ou pendjabi occidental; puis, à partir du confluent de l'Indus et du Panjnad, le *sindhi*, dont le représentant le plus méridional, le *larou*, a servi de langue littéraire. En suivant la côte on rencontre ensuite le *guzrati*, qui a une littérature depuis le XVe s.; puis le *marathe*, qui s'étend jusqu'aux confins méridionaux du domaine indo-aryen [1]. Les plus anciens documents épigraphiques du marathe sont du début du XIIe s. de notre ère; il y a en marathe une littérature poétique assez riche, fondée actuellement sur le parler du deç (entre les Ghats et la frontière du Berar). En remontant de là à l'Est vers le bassin du Gange on rencontre une série de dialectes intermédiaires entre le guzrati et l'hindoustani: ce sont d'abord les dialectes *bhil* (parlés par des tribus de race munda), puis les dialectes *rajasthani*, dont le principal est le *marvari*, au N.-O. du domaine.

Le groupe central comprend : le *penjabi* proprement dit, qui est notamment la langue des Sikhs et par suite a été transporté dans d'autres régions de l'Inde et jusqu'en Chine; le *pahari*, le plus septentrional et qui touche au domaine tibéto-birman; l'*hindi occidental* (entre la frontière du Pendjab et Cawnpore), l'*hindi oriental* (de Cawnpore à peu près jusqu'à Bénarès), et le *bihari*. Ces quatre derniers comportent de nombreux dialectes; ainsi l'hindi oriental a comme principal dialecte l'*awadhi* parlé surtout entre Lakhnau, Allahabad et Aoude; le bihari comprend le *bhojpuri*, le *magahi* et au delà du Gange le *maithili*. Ce dernier a un passé littéraire.

L'hindi occidental est important parce que son principal dialecte est l'*hindoustani*. Le berceau de l'hindoustani est aux environs et au Nord de Mirat, mais l'hindoustani a pris la valeur d'une langue commune dans le bazar attaché à la cour de Delhi et de là s'est étendu dans l'Inde du Nord. Il comporte plusieurs variétés : il a une forme littéraire, *l'ourdou*, qui s'écrit avec

1. J. Bloch, *La formation de la langue marathe*, Paris, 1914-1919, p. 8 et ss.

l'alphabet arabe et use largement du vocabulaire persan ; l'ourdou est employé spécialement dans l'Hindoustan occidental par les Musulmans et les Hindous qui ont subi l'influence persane. Une autre forme littéraire de l'hindoustani porte simplement le nom d'*hindi* ; de création récente comme tel, l'hindi, noté en alphabet indien, représente une réaction contre l'ourdou et a été plus ou moins artificiellement imprégné de mots sanskrits.

Le groupe oriental comprend le *bengali*, l'*oriya* et l'*assamais*, qui sont extrêmement semblables, et qui ont tous trois une littérature ancienne ; la plus riche et la plus belle est toutefois celle du bengali. Mais le bengali littéraire est chargé d'emprunts à l'hindi et au sanskrit. Le domaine de l'assamais est dans la vallée du Brahmapoutra depuis la région à l'Est de Koutch-Béhar jusqu'à Dibroughar (en dehors de la planche).

Dans la partie méridionale de l'île de Ceylan, s'emploie encore un parler indo-aryen, le *singhalais* [1].

Enfin, en dehors de l'Inde, l'indo-aryen a un rameau détaché dans le *tsigane*, qui a été transporté par des migrations de nomades à travers la Perse et l'Arménie jusqu'aux régions de l'Europe occidentale. Le tsigane appartient originellement au groupe indien du Nord-Ouest ; il s'en est détaché à partir du ve s. de notre ère [2], mais il a subi de nombreuses transformations suivant les langues avec lesquelles il s'est trouvé en contact, et il se présente parfois avec un caractère de langue secrète, dont le vocabulaire seul est tsigane, mais dont la grammaire est empruntée à la langue du pays ; c'est le cas pour la langue des Tsiganes d'Arménie [3].

1. W. Geiger, *Litteratur und Sprache der Singhalesen* dans le *Grundriss der indo-arischen Philologie*, Strassburg, 1900.

2. De Goeje, *Mémoire sur les migrations des Tsiganes à travers l'Asie*, Leide, 1903 ; F. Miklosich, *Ueber die Mundarten und die Wanderungen der Zigeuner Europa's* (*Denkschriften* de l'Académie de Vienne, vol. XXVI à XXXI, 1877-1880).

3. N. Finck, *Die Sprache der armenischen Zigeuner* (Mém. de l'Académie impér. de Pétersbourg, VIIIe série, 1907). Sur le tsigane d'Espagne, langue des « Gitanes », v. Keller, *Zeitschrift für rom. Phil.* XVI, 165.

Deux types d'écriture ont servi, depuis le IIIe s. avant notre ère, à noter les langues indo-aryennes. Ils sont probablement tous deux d'origine sémitique, mais la formation n'en est pas connue avec précision. L'écriture la plus répandue est l'écriture *brāhmī*, qui comporte d'ailleurs certaines variétés d'alphabets et dont la forme demeurée usuelle porte le nom de *devanāgarī* ; on rencontre des alphabets dérivés de la brāhmī sur toute l'étendue du domaine indo-aryen, du Nord au Sud et jusqu'en Indochine. L'autre type d'écriture, exclusivement limité au Nord-Ouest du domaine et qui n'a pas survécu, est l'écriture *kharoṣṭhī*. La kharoṣṭhī s'écrit de droite à gauche, la brāhmî de gauche à droite. Toutes deux notent les voyelles en même temps que les consonnes [1].

*
* *

GROUPE IRANIEN [2].

Le groupe iranien ne le cède guère au groupe indien en étendue et en importance. Il a produit une langue religieuse, le zend, et deux grandes langues communes, le sogdien et le pehlvi, devenu ensuite le persan, qui ont servi de langues de civilisation à une partie de l'Asie pendant des siècles. Mais l'histoire de l'iranien est difficile à suivre ; sur la période ancienne, où la langue a évolué avec une grande rapidité, on n'a que des documents isolés, offrant rarement une continuité ; en outre, un bon nombre des dialectes iraniens ne sont connus qu'à l'époque moderne et par des renseignements fragmentaires et misérables.

A la date la plus ancienne, l'iranien apparaît sous deux formes, qui sont toutes deux de l'iranien occidental, le vieux-perse et le zend.

Le VIEUX-PERSE [3] est le dialecte de la Persis, à l'extrémité sud-ouest du domaine iranien. Il n'est connu que par les inscriptions

1. Voir Bühler, *Indische Paläographie* (avec 17 planches), dans le *Grundriss der indo-arischen Philologie*, Strassburg, 1896 ; J. Wackernagel, *Altindische Grammatik*, t. I, p. lvj et suiv.

2. Consulter la carte, p. 43.

3. A. Meillet, *Grammaire du vieux-perse*, Paris, 1915.

en caractères cunéiformes des rois achéménides Darius (522-486) et Xerxès (486-466). Les inscriptions postérieures sont rares et peu intéressantes. Il ne s'agit pas d'une langue littéraire antérieurement fixée, ni d'une langue commune, ni même à proprement parler d'une langue de bureaux de chancellerie. Il s'agit d'un dialecte qui était le parler maternel des souverains et de leur entourage, et dans lequel ils ont tenu à faire transcrire, sous une forme manifestement sincère, les textes qui établissaient leur puissance et célébraient leurs exploits. Les inscriptions des Achéménides sont bilingues ou même trilingues : outre le vieux-perse, langue de l'aristocratie conquérante, elles sont écrites en élamite, langue du royaume de Cyrus, le fondateur de la dynastie, et en babylonien, vieille langue de civilisation asiatique. On les trouve disséminées dans des parties diverses de l'empire : à Persépolis, à Nakhš-i-Roustem (tombeau de Darius), à Suez, à Elvend. La plus grande de toutes, qui est trilingue, est gravée sur le flanc d'une montagne à Béhistoun, en Médie, à une lieue environ au nord-est de Kermanchah.

Le ZEND OU AVESTIQUE est la langue d'un texte religieux, l'Avesta, malaisée à fixer avec certitude aussi bien dans l'espace que dans le temps. La rédaction définitive de cette bible du mazdéisme, telle qu'elle nous est parvenue, ne remonte pas plus haut que l'époque des Sassanides. Ceux-ci, afin de mieux affermir la religion d'état qu'ils avaient fondée, firent constituer un texte liturgique avec des morceaux de dates diverses, réunis tant bien que mal par des raccords et complétés par des ajoutures. On y rencontre deux séries de textes, différents par l'orthographe et la grammaire, et d'étendue fort inégale : d'une part les *Gāthās* (« chants »), d'un archaïsme à peu près égal, parfois supérieur, à celui du Rig-Veda ; d'autre part, l'*Avesta* proprement dit, vaste compilation dont la langue n'a pas une parfaite unité. Le texte de l'Avesta a d'abord été écrit dans un alphabet sémitique, de forme araméenne, et il a été ensuite transcrit, par des gens qui n'avaient pas de tradition certaine, dans l'écriture spéciale qui nous a été conservée. La lecture exige donc constamment une interprétation, parfois des moins sûres. Un examen minutieux du zend démontre tou-

tefois que cette langue représente une forme dialectale peu éloignée du vieux-perse. C'est la Médie que la légende assigne comme patrie à Zoroastre, le fondateur du mazdéisme. On doit considérer le zend comme issu d'un dialecte, septentrional par rapport au vieux-perse, mais uni assez étroitement à lui et appartenant comme lui au groupe iranien occidental.

En dehors du vieux-perse et du zend, on ne possède aucun document d'un parler iranien de forme ancienne. Cela est fort regrettable ; car les rares mots ou noms propres que les Grecs nous ont transmis de la langue des Scythes révèlent un état linguistique très différent de celui que présente la langue des Perses et des Mèdes ; le SCYTHE appartenait en effet au groupe iranien septentrional : c'était proprement de l'iranien du Nord-Ouest (v. p. 42).

Plusieurs siècles après l'ère chrétienne, l'iranien reparaît dans l'histoire sous une forme très évoluée, qu'on désigne en général du nom de moyen-iranien. Le représentant le mieux connu en est le PEHLVI [1], qui appartient au groupe perse et est l'ancêtre direct du persan moderne. Il n'y a entre le vieux-perse, le pehlvi et le persan d'autres différences que celles qui résultent du développement linguistique. Bien que le vieux-perse présente quelques traits dialectaux qui ne se retrouvent pas en pehlvi, c'est au fond la même langue. Entre le pehlvi et le persan, il n'y a presque pas de solution de continuité : les textes du persan ancien peuvent être indifféremment baptisés aussi pehlvi récent. Les Arsacides (de 250 av. J.-C. à 226 ap. J.-C.) n'ont laissé aucune inscription en pehlvi (voir page 37) ; mais on en a du fondateur de la dynastie sassanide, Artachir (226-241). Le pehlvi a été la langue officielle de l'État et de l'Église pendant toute la durée de l'empire sassanide (226-652). L'écriture en est d'origine sémitique (araméenne) ; l'influence sémitique se fait sentir également sur la langue.

1. Voir E. Blochet, *Études de grammaire pehlvie*, Paris (sans date), et surtout Salemann, *Grundr. der iran. Phil.*, I, 1, p. 249-332.

Le pehlvi est connu surtout par des textes mazdéens conservés chez les Zoroastriens de la Perse et de l'Inde et aussi par des textes manichéens récemment découverts en Asie Centrale. Ces derniers qui remontent en partie à Mani lui-même (IIIe s. de notre ère) sont écrits en une langue plus conforme à l'usage parlé que le pehlvi des Sassanides ; et l'on n'y rencontre pas les équivalents sémitiques qui, dans la notation des inscriptions sassanides et du pehlvi de l'Église mazdéenne, masquent la plupart des mots usuels. Ils ont fait faire à la connaissance du pehlvi un grand progrès.

Au moyen-iranien appartiennent encore les documents écrits en pehlvi du Nord et en sogdien.

Le nom de pehlvi du Nord s'applique à deux séries de textes : d'une part des textes manuscrits, de caractère manichéen, mais présentant une forme dialectale différente de ceux qui viennent d'être mentionnés ; d'autre part des textes épigraphiques figurant à côté du pehlvi officiel sur les inscriptions sassanides. La langue de ces derniers, qu'on appelle parfois à tort *chaldéo-pehlvi*, n'est qu'une forme du PARTHE, langue de la dynastie arsacide, qui ne l'a malheureusement pas utilisée dans son épigraphie officielle (rédigée en araméen ou en grec). C'est par le parthe que l'iranien a agi sur l'arménien (v. p. 46).

Le SOGDIEN [1] n'est connu que depuis les découvertes faites en Asie Centrale dans les vingt dernières années. Les textes déchiffrés (quelques-uns chrétiens, mais la plupart bouddhiques) sont dans une écriture d'origine araméenne, différente du pehlvi. Ils appartiennent en majorité au VIIIe-IXe s. de notre ère ; mais on en a trouvé de plus anciens, remontant même au début de l'ère chrétienne. Ils attestent le grand rôle qu'a joué le sogdien, comme langue de civilisation, depuis la Sogdiane proprement dite jusqu'en Mandchourie, où l'inscription trilingue de Kara-Balgassoun (IXe s. ap. J.-C.) a une partie en sogdien [2]. Le sogdien, qui

1. R. Gauthiot, *Essai sur le vocalisme du sogdien*, Paris, 1913 ; premier fragment de la *Grammaire sogdienne*, dont la première partie (Phonétique) a paru en 1923, et dont la fin est en préparation (par les soins de M. Benveniste).

2. F. W. K. Müller, *Sitzber. der Berl. Akad. d. Wiss.*, 1909, p. 276.

comporte d'ailleurs quelques variétés dialectales, est de l'iranien du Nord-Est.

Les découvertes faites au Turkestan ont révélé une langue écrite du groupe iranien, employée par les bouddhistes de la région de Khotan ; on en a divers textes nombreux, encore inédits et non étudiés pour la plupart. Cette langue, qui est sans doute celle des *Saka,* est souvent désignée d'une manière vague sous le nom d' « iranien oriental ».

Parmi les langues iraniennes parlées aujourd'hui, il faut donner une place d'honneur au PERSAN [1], grande langue commune représentant une belle civilisation et dotée d'une riche littérature. Les premiers documents en sont du IXe s. et se distinguent à peine du pehlvi du Sud (v. p. 36). Dès cette époque la langue est fixée ; elle n'a que peu changé depuis. Elle s'est imposée aux écrivains, même quand ils étaient originaires, comme c'est le cas du grand poète Firdousi (Xe s.), d'une région de l'Iran dialectalement différente. Le persan est noté au moyen de l'écriture arabe complétée par quelques signes diacritiques. Le persan de Firdousi a un vocabulaire tout iranien ; mais, avec le temps, le vocabulaire du persan écrit s'est chargé de plus en plus d'emprunts à l'arabe, si bien qu'il est devenu plus arabe qu'iranien. Au cours du moyen âge, le persan a pris comme langue commune une grande extension vers l'Est ; il a été parlé jusque dans l'Inde, par exemple à la cour de Delhi, et on en trouve la trace dans l'ourdou (v. p. 32-33). En dehors de la Perse, le persan est aujourd'hui parlé dans des parties notables de l'Afghanistan et du Turkestan.

En Perse même, le persan n'est pas la seule langue en usage. Il recouvre comme langue commune un certain nombre de parlers locaux plus ou moins bien connus, dont quelques-uns semblent des restes de parlers iraniens distincts. Tels sont par exemple au centre de la Perse [2] les dialectes de Kachan, de

1. P. Horn, *Grundriss der iranischen Philologie,* Strassburg, I, 2te Abt., p. 1 et ss.

2. W. Geiger, *Grundriss der iran. Philol.,* t. I, 2te Abt., p. 381.

Nayin [1], de Chiraz, de Sivend [2] (situés respectivement au N., à l'E. et au S. d'Ispahan). Tel est surtout le dialecte des Parsis ou Guèbres, parlé dans certaines villes (notamment à Yezd) par une dizaine de mille d'individus ; on le désigne parfois à tort sous le nom de *deri* [3]. C'est dans l'Inde (présidence de Bombay) que les Parsis ont leurs communautés les plus nombreuses et prospères [4] ; mais les Parsis de l'Inde parlent guzrati.

Enfin, le persan se trouve en Perse même en concurrence avec d'autres langues, comme le turc, l'arabe, l'arménien, l'afghan, le beloutchi (p. 40), ainsi qu'avec les dialectes caspiens et le kurde.

Le principal des DIALECTES CASPIENS [5] est le *mazandarani* dont le nom sort du *mazainya daēva* de l'Avesta ; il est connu comme langue littéraire dès le moyen âge et a été employé par des poètes restés populaires. A ce dialecte se rattache étroitement le *semnâni*, parlé autour de Semnan (au S.-E. du précédent) ; plus ou moins différents sont le *ghilaki*, dans la région de Recht, et le *tâliche*, dont le domaine s'étend jusqu'au fleuve Aras, dans l'Azerbeidjan. On trouve même un dialecte caspien, le *tât*, plus au Nord encore, dans la presqu'île d'Apchéron, près de Bakou. Ces dialectes caspiens forment un groupe à part, très différent du persan, et ayant des traits communs avec l'iranien septentrional (v. p. 43) [6].

Le domaine du KURDE [7] est difficile à délimiter, parce que cette langue est parlée par des tribus en partie nomades qui l'ont transportée vers le Nord jusqu'aux environs d'Erivan et de Kars en Arménie russe, vers l'Ouest jusqu'en Cilicie et même en Syrie, vers l'Est dans la Perse centrale, le Khorassan et jusqu'en Afghanistan. Le centre du domaine kurde est la région monta-

1. Querry, *M. S. L.*, IX, p. 110-124.
2. Cl. Huart, *Journal asiatique*, 1889, t. II, p. 241-265.
3. Cl. Huart, *ibid.*, 1888, t. I, p. 298-302.
4. D. Menant, *Les Parsis*, Paris, 1898, p. 54.
5. W. Geiger, *Grundr. der iran Phil.*, t. I, 2te Abt., p. 344.
6. Gauthiot, *M.S.L.*, XIX, 135.
7. A. Socin, *Grundr. der iran. Phil.*, I, 2te Abt., p. 249.

gneuse qui sépare la Perse de la Turquie (monts Zagros) ; il comprend surtout le Kurdistan en territoire turc. On y peut distinguer trois groupes dialectaux : le kurde occidental ou kurmandji, parlé dans les régions de Diarbékir, Mardin, Saert et Djoulamerk, mais qui s'étend beaucoup plus loin vers l'Ouest ; le kurde oriental parlé dans les régions de Souleimanieh et Senna ; le kurde méridional parlé dans la région de Kermanchah et plus au Sud dans le pays des Bakhtiares et des Lauris. Ces trois groupes comprennent eux-mêmes des parlers assez différents. Le kurde appartient à l'iranien occidental, mais il a quelques traits communs avec l'iranien septentrional [1]. La littérature kurde, presque uniquement orale, comprend des fragments d'épopée, des contes (d'inspiration persane) et surtout des chants populaires.

L'iranien oriental est aujourd'hui représenté surtout par le beloutchi, l'afghan et les dialectes du Pamir. Il n'a aucune unité.

Le domaine du BELOUTCHI [2] est limité approximativement par la mer d'Oman, le cours de l'Indus jusqu'à Dera-Ghazi-Khan, le désert de Registan et le cours du Hilmend, le plateau de Sarhad et le Mekran, y compris Bampour en territoire persan. Le centre de ce domaine est occupé par les Brahouis, qui parlent une langue dravidienne. Il y a donc dans le beloutchi deux groupes dialectaux séparés : un beloutchi septentrional, qui porte dans son vocabulaire la marque d'influences indiennes, et un beloutchi méridional dit aussi *mekrani*. Ils sont très différents l'un de l'autre et se divisent d'ailleurs en de nombreux sous-dialectes. Le nom du beloutchi apparaît pour la première fois dans le Chah-nameh de Firdousi ; la littérature en est pauvre et toute orale : elle ne comprend guère que des contes (inspirés du persan) et des chants populaires (beaucoup moins importants que ceux des Afghans).

L'AFGHAN [3] ou *puštū* est la langue des Afghans, qui se

1. Gauthiot, *M. S. L.*, XIX, 135.

2. W. Geiger, *Grundr. Iran Phil.* II, 2te Abt., p. 231.

3. W. Geiger, *Grundr. Iran. Phil.* I, 2te Abt., p. 201 ; J. Darmesteter, *Chants populaires des Afghans*, Paris, 1888-1890.

nomment eux-mêmes Puštāna (singul. Puštūn) ; il a une autre importance que le beloutchi. Sa littérature écrite commence au XVIe s. et il a connu au siècle suivant une floraison poétique, inspirée en grande partie de modèles persans. Il possède aujourd'hui une littérature orale des plus riches, surtout des contes et des chansons populaires. Les limites de l'afghan ne coïncident pas avec celles de l'émirat d'Afghanistan. Dans quelques parties du territoire de l'émirat on parle persan, beloutchi, mongol, turc, arabe, etc. ; en revanche, on parle afghan sur quelques points du territoire persan et surtout dans l'Inde anglaise le long du cours de l'Indus. On compte en tout environ 4 millions d'Afghans, dont 1 million 1/2 dans l'Inde anglaise. Il y a un afghan commun qui se parle partout à peu près de même ; mais on compte un certain nombre de parlers locaux variés, qui se ramènent à deux groupes de dialectes, l'un septentrional (Kaboul, Pechaver), l'autre s'étendant au Sud et à l'Ouest.

Dans la région du Pamir sont parlés aujourd'hui encore de nombreux dialectes iraniens, survivants d'une grande expansion iranienne de date fort ancienne. On a trouvé en effet en Asie Centrale des textes bouddhiques d'alphabet indien mais de langue iranienne, provenant des vallées septentrionales de l'Hindou-Kouch et même de la région de Khotan (v. p. 38). Aujourd'hui, dans les vallées du Pamir, les groupes parlant iranien sont noyés au milieu de populations parlant soit des dialectes indiens, soit des dialectes turcs. Ils se donnent le nom commun de *ghaltcha* et se répartissent linguistiquement en plusieurs parlers qui ont de nombreux traits communs [1] : le *wakhi*, l'*ichkâchimi*, le *chighni*, le *rochani*, échelonnés le long du cours du Pandj, fleuve qui contribue à former l'Amou-Darya ; le *sariqoli*, parlé sur le versant opposé (E.) du Pamir ; le *sanglitši*, et le *mindžani* [2] parlés dans des vallées au N. de l'Hindou-Kouch ; le *yidgha* parlé dans une vallée au S. de l'Hindou-Kouch. Ces derniers se rapprochent par certains traits des dialectes caspiens et sont peut-être les

1. W. Geiger, *Grundr. Iran. Phil.*, I, 2te Abt., p. 288.
2. R. Gauthiot, *M.S.L.*, XIX, p. 133-157.

restes d'un groupe iranien central, distinct à la fois du groupe perse, du groupe scythique et du groupe afghan.

Au Nord-Ouest du Pamir, dans la vallée du Yagnob, affluent du Zarafchan, on rencontre enfin un dialecte iranien, qui paraît appartenir à un groupe différent des précédents : c'est le *yagnobi* [1], qui descend de l'ancien sogdien et représente un parler iranien septentrional.

Un autre représentant de l'iranien septentrional apparaît fort loin de là dans l'OSSÈTE [2], misérable reste du vaste ensemble de parlers scythes qui ont dominé jadis dans la Russie méridionale avant que le slave les en ait chassés. Le domaine de l'ossète est dans un coin du Caucase à l'O. de Vladikaucase (voir pl. 8); il forme, à une exception près (v. ci-dessous), un groupe linguistique compact, entouré de toutes parts par des langues d'un tout autre type. On en distingue deux variétés : l'ossète oriental ou tagaurique, le plus important, parlé dans la vallée des fleuves Ardon, Sadon et Giseldon par les Allaghirs, les Kurtates et les Tagaures, avec un sous-dialecte, le toualique, parlé au Sud par les Touales (oss. Tvaltä) ; l'ossète occidental ou digorique, parlé dans le bassin du fleuve Ouroukh par les Digores, avec un petit noyau détaché dans le district de Mozdok, sur le Térek. Il n'y a guère plus de 17.000 individus parlant ossète, dont 5.000 dans le district de Mozdok. On a imprimé depuis trente ans quelques textes ossètes; c'est le dialecte oriental qui a servi à la traduction de la Bible dans la première moitié du XIXe siècle.

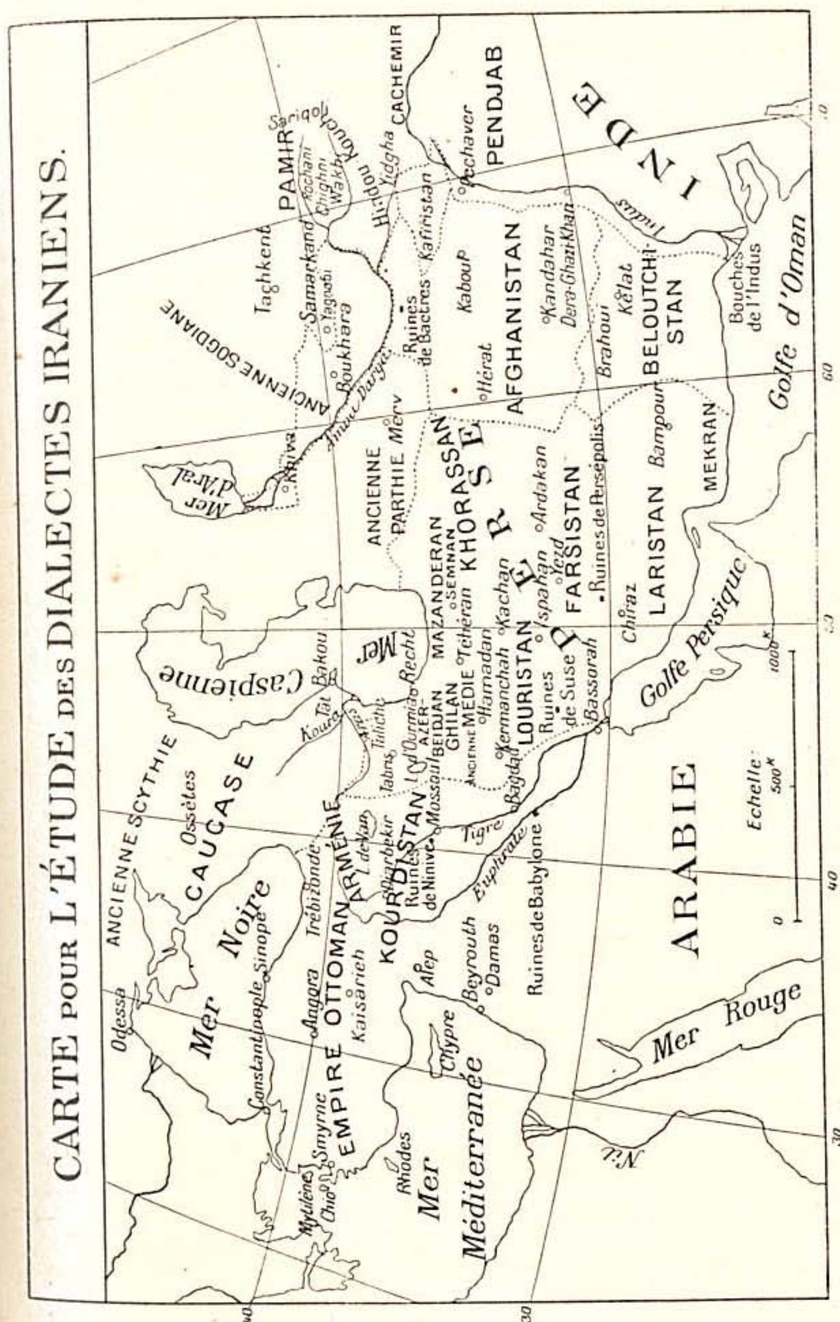
*
* *

Parmi les langues indo-européennes de l'Asie, il faut ranger aussi celle qu'on appelle le TOKHARIEN, qui constitue un groupe indépendant [3].

1. R. Gauthiot, *Comptes-rendus de l'Académie des Inscriptions et Belles-lettres*, 1913, p. 671 et ss.

2. Wsewolod Miller, *Grundr. der iran. Phil.*, Anhang zum Iten Bande.

3. Voir l'article de A. Meillet, *Indogermanisches Jahrbuch*, t. I, p. 1-19, où



CARTE POUR L'ÉTUDE DES DIALECTES IRANIENS.

Ce nom a été emprunté à l'historien Strabon, qui mentionne un peuple de *Tokharoi* (XI, p. 511), appelé *Tukhāra* en indien et *Tu-hu-la* en chinois, mais sans que l'on sache rien de précis sur lui. On a appliqué, peut-être un peu légèrement, ce nom de tokharien à une langue dont quelques textes, fragmentaires en général, ont été recueillis au début du xxe siècle en Asie Centrale (Turkestan chinois). Ils sont notés en une écriture d'origine indienne et s'inspirent fortement de la littérature sanskrite, ce qui en a facilité le déchiffrement. Ce sont des fragments de textes médicaux et surtout des textes religieux bouddhiques, en partie traduits du sanskrit. Le « tokharien » n'était pas seulement une langue savante et religieuse ; on le parlait comme langue vivante dans le pays dont les documents proviennent. Ces documents révèlent l'existence de deux dialectes différents, que l'on a d'abord distingués en tokharien A et tokharien B, mais qu'on baptise aussi d'après la région où les documents ont été trouvés, l'un, le tokharien de Tourfan ou tourfanien, l'autre le tokharien de Koutcha ou koutchéen. Parmi les documents en koutchéen se trouvent des laissez-passer pour caravanes, portant le nom d'un roi Suvarnate, qui régnait d'après les témoignages chinois dans la première moitié du viie siècle de notre ère. Cela nous indique qu'à cette date le tokharien B était d'un emploi courant dans la région. On ignore la date à laquelle les deux dialectes se sont éteints.

*
* *

Arménien.

Le groupe arménien est localisé dans le grand pays montagneux qui s'étend entre la Mésopotamie, les vallées méridionales du Caucase et la côte sud-orientale de la mer Noire [1].

est résumée l'histoire du déchiffrement (auquel ont particulièrement contribué MM. A. Meillet et S. Lévi pour le koutchéen) et indiquée la place du « tokharien » dans la famille indo-européenne.

1. Voir A. Meillet, *Esquisse d'une grammaire comparée de l'arménien classique*. Vienne, 1903, et *Altarmenisches Elementarbuch*, Heidelberg, 1913 ; H. Adjarian, *Classification des dialectes arméniens*, Paris, Champion, 1909 (avec une carte).

Le nom par lequel les Arméniens se désignent est celui de *Hay*, pl. *Hayk'*. Mais les inscriptions achéménides (v. p. 35) nomment déjà les Arméniens du nom d'*Armina-*, *Arminiya-*, que nous leur donnons encore. Il semble que ce soit entre le VIIIe et le VIe s. avant notre ère que l'arménien a été apporté dans la région qui porte historiquement le nom d'Arménie. Cette région était antérieurement occupée par les peuples qui ont laissé les inscriptions vanniques et les inscriptions dites des Mitani [1].

Les plus anciens manuscrits de l'arménien sont ceux d'une traduction de l'Évangile ; ils ne remontent pas plus haut que le IXe siècle de notre ère. Mais la langue dans laquelle cette traduction fut faite avait été fixée, d'après la tradition, au début du Ve s., et en tout cas, au plus tard vers la fin du VIe s. ou le début du VIIe s. ; et l'écriture arménienne aurait été inventée à cette époque par un homme d'église, nommé Mištoc, précisément afin de faciliter la traduction des Saintes Écritures et de rendre ainsi l'Arménie indépendante des églises grecque et syriaque. L'écriture arménienne, qui compte 36 signes en partie imités du grec, est un chef-d'œuvre d'exactitude phonétique.

On appelle vieil-arménien la langue dans laquelle la traduction de la Bible a été faite. C'est la langue classique, dite en arménien *grabar* (« langue écrite »), dont la forme la plus ancienne se trouve dans la traduction de l'Évangile et dans divers ouvrages, notamment dans les écrits de l'évêque Eznik. Le dialecte qui servit de base à cette langue classique était parlé, semble-t-il, dans la région de Tarawn, sur les bords du lac de Van. Au cours du moyen âge, il s'est développé une littérature arménienne assez abondante, surtout historique et théologique.

Aujourd'hui, l'arménien se parle sur une aire étendue, non seulement sur le petit territoire de la république d'Arménie (autour d'Erivan), mais dans la république d'Azerbeidjan, dans plusieurs provinces de Turquie (région de Mouch, Van, Bitlis, etc.), et dans un coin de la Perse. De plus la langue a été transportée au delà des frontières de l'ancien royaume d'Arménie. Il y a des colonies arméniennes abondantes et prospères en Géorgie

1. Voir Fossey, dans *Scientia*, août 1922, p. 97.

et dans les provinces russes du Don, en Perse, en Asie Mineure (Alexandrette, Smyrne), en Turquie d'Europe (où l'on compte environ 300.000 Arméniens à Constantinople, Andrinople, Rodosto, etc.), en Bulgarie, en Roumanie (notamment dans les territoires de l'Ardéal, à Elisabetpol, Arménopol et à Suczawa), aux Indes, en Égypte et enfin aux États-Unis d'Amérique. Le total des individus parlant arménien pouvait être de trois à quatre millions en 1914.

On peut répartir les parlers arméniens modernes en deux groupes principaux, l'un occidental (dit aussi arménien de Turquie), comprenant les parlers d'Erzeroum, de Mouch, de Van, de Diarbékir, d'Akn. de Sivas, etc., l'autre oriental (dit aussi arménien de Russie) comprenant les parlers d'Erivan, de Tiflis, du Karabagh et de la côte occidentale de la Caspienne. C'est au groupe occidental qu'appartiennent les parlers des colonies arméniennes du bassin du Don, d'Asie Mineure et de Constantinople. A l'intérieur des deux groupes se rangent des parlers assez dissemblables, quelques-uns même nettement aberrants, comme le parler d'Agulis à l'O. du Karabagh ou d'Artvin à l'E. de Trébizonde, surtout comme ceux de la région entourant le lac d'Ourmia (Xoy, Ourmia, Maragha), que l'on considère parfois comme formant un groupe dialectal à part des deux autres.

*
* *

Il n'existe sur le passé de l'arménien antérieurement au ve s. aucun document utilisable. L'influence iranienne s'est fortement exercée sur le vocabulaire, par suite du fait que de 66 à 387 ap. J.-C. le pays a été dominé par une aristocratie parthe. Lorsque s'est constituée la grammaire comparée des langues indo-européennes, on a même cru d'abord qu'il fallait rattacher l'arménien au groupe iranien. Mais le linguiste allemand H. Hübschmann a démontré que l'arménien formait un groupe indépendant de tous les autres [1]. En fait, on ne connaît même pas à

1. En 1873, dans un article de la *Zeitschrift für vergleichende Sprachforschung*, t. XXIII, p. 5-42.

l'arménien de groupe voisin, qui lui soit étroitement apparenté comme le baltique l'est au slave ou le latin au celtique. Dans la famille indo-européenne, l'arménien est isolé.

*
* *

Peut-être cet isolement cesserait-il si nous connaissions mieux les langues de deux peuples que l'histoire ancienne mentionne au Nord du monde hellénique, les Thraco-Phrygiens et les Macédoniens. Les Thraces, au dire d'Hérodote (V, 3), ne le cédaient en nombre parmi les peuples de la terre qu'aux seuls Indiens. Ils avaient installé de nombreuses colonies en Asie Mineure (Strabon XIII, p. 586). Ils étaient apparentés aux Gètes, dont la langue était la même que celle des Daces (Strabon VI, p. 305). Les Phrygiens habitaient auprès des Macédoniens avant de passer en Asie Mineure, où les Arméniens n'étaient que leur colonie (Hérodote, VII, 73). On entrevoit donc l'existence d'un groupe thraco-phrygien, qui n'est malheureusement représenté pour nous que par de rares documents linguistiques : gloses et noms propres en THRACE, gloses, noms propres et quelques inscriptions en PHRYGIEN [1]. Il n'est pas douteux cependant qu'il s'agisse d'une langue indo-européenne.

Le MACÉDONIEN n'est pas mieux connu ; les documents qu'on en possède ne permettent pas de décider s'il faut le considérer comme un groupe à part ou le rattacher à un des groupes existants. Certains linguistes sont enclins à y voir une forme particulière et en tout cas assez aberrante du groupe hellénique [2].

1. Sur le thraco-phrygien, voir P. Kretschmer, *Einleitung in die Geschichte der griechischen Sprache*, p. 217 ; H. Hirt, *Die Indogermanen*, II, p. 592. Sur le thrace en particulier, voir Tomaschek, *Die alten Thräker* dans les *Sitzungsberichte* de l'Académie de Vienne, tomes 128, 130 et 131 (1893-1894) ; à consulter aussi les travaux de M. G. Seure dans le *Bulletin de Correspondance hellénique* et dans la *Revue des Études anciennes*. M. Seure prépare un dictionnaire de la langue thrace.

2. Sur le macédonien, voir, outre les ouvrages déjà cités de P. Kretschmer, p. 283, et de H. Hirt, p. 602, le livre de O. Hoffmann, *Die Makedonen, ihre Sprache und Volkstum*, 1906, et un article de V. Lesny, *Kuhn's Zeitschrift*, XLII, p. 297.

Groupe hellénique.

Le groupe hellénique se résume en ce qu'on désigne du nom général de *grec*.

Les régions dans lesquelles devait fleurir la langue grecque étaient habitées avant l'arrivée des Grecs par des populations de race et de langue inconnues dont nous n'avons guère conservé que les noms : Pélasges, Lélèges, Dryopes, Cariens, etc. [1]. Le *pélasge* était encore parlé, semble-t-il, au v[e] s. avant notre ère, sur la côte de Thrace, au Sud de la Propontide et dans certaines îles comme Imbros, comme Lemnos (où une fameuse inscription découverte en 1885 en fournit peut-être un document, d'ailleurs non interprété). Certains détails de structure rendent vraisemblable la possibilité d'une parenté entre le pélasge et l'étrusque. Il y eut des Pélasges même en Crète. Un autre peuple de cette île, les Étéocrétois, parlait sans doute une langue qui n'était pas indo-européenne. Le substrat sur lequel s'est étendu et développé le grec était donc des plus mêlés ; cela explique en partie que la structure grammaticale et le vocabulaire du grec présentent tant d'innovations qui lui sont propres [2].

La plus ancienne inscription grecque qui soit datée est sans doute une de celles qu'on a trouvées à Abou Simbel (en Égypte, sur le Nil, aux confins de la Nubie) ; elles se rattachent à l'expédition de Psammétique II, roi d'Égypte, contre l'Éthiopie, en 591 avant notre ère. Mais il y a sur des vases attiques des inscriptions qui paraissent remonter au VII[e] et même au VIII[e] siècle. A partir du v[e] s. le nombre des inscriptions grecques va en augmentant et s'étend à tout le monde hellénique. C'est pour nous un ensemble de témoignages des plus précieux pour fixer l'histoire

1. P. Kretschmer, *Einleitung in die Geschichte der griechischen Sprache*. Göttingen, 1896.

2. A. Meillet, *Aperçu d'une histoire de la langue grecque*, Paris, Hachette, 2[e] édition 1921 ; O. Hoffmann, *Geschichte der griechischen Sprache*, Berlin-Leipzig, coll. Göschen, 2[e] édit. 1921 ; Zarncke, *die Entstehung der griechischen Literatursprachen*, Leipzig, 1890.

de la langue. Il complète et précise l'enseignement que l'on tire des textes littéraires, dont les manuscrits sont de date bien postérieure : les plus anciens manuscrits ne remontent guère en effet plus haut que le IIIe ou IVe siècle de notre ère [1]. Le sol de l'Égypte a, il est vrai, livré des papyrus qui sont du IIIe et même du IVe av. J.-C., mais les papyrus ne contiennent guère d'œuvres littéraires complètes et suivies.

A l'époque des plus anciens documents, le grec apparaît déjà sous des formes variées représentant autant de dialectes différents. Il se pose donc à propos du grec ancien un délicat problème de répartition dialectale. Malheureusement ce problème n'admet pas de solution absolument précise. La formation et l'extension des dialectes grecs résultent de faits historiques particuliers, qui ont déterminé la façon dont les populations se sont groupées et ont agi les unes sur les autres. Or ces groupements et ces actions sont loin d'être simples. Les migrations ont été successives et échelonnées ; des groupes déjà différenciés y ont d'ailleurs pris part. Souvent les nouveau-venus ont balayé les envahisseurs antérieurs, ne laissant subsister leurs établissements que dans quelques îlots. Les invasions nouvelles eurent donc souvent pour résultat des exodes de populations et des fondations de colonies outre-mer. Presque aucun des peuples de la Grèce n'était venu directement à l'endroit où nous les trouvons installés à date historique ; la plupart avaient au préalable fait des stations sur d'autres points, d'où ils avaient été délogés par la pression de nouveaux arrivants. Aussi y a-t-il dans le monde hellénique un grand enchevêtrement de dialectes et presque pas de dialectes purs. Les plus purs sont dans des îlots, restés en dehors des grands courants d'immigration ; ainsi l'arca-

1. Voir Maunde Thompson, *Greek and Latin Paleography*, Oxford, 1912. Les plus anciens manuscrits grecs sur parchemin, écrits en onciale, sont l'Ambrosianus de l'Iliade (peut-être du IIIe s. ap. J.-C., d'après Thompson) et trois manuscrits de la Bible, le Vaticanus (IVe s.), le Sinaïticus (fin du IVe s.), l'Alexandrinus (Ve s.). L'usage du parchemin est d'ailleurs beaucoup plus ancien ; on a trouvé des fragments de manuscrits sur parchemin qui paraissent remonter au Ier siècle de notre ère (Thompson, *Paleogr.*, p. 30).

dien confiné au centre du Péloponnèse ou le lesbien de certaines îles.

On peut cependant répartir en quelques grands groupes les dialectes de la Grèce ancienne [1] :

a) le groupe *ionien-attique*, le plus important au point de vue littéraire. L'Ionie a été en effet la première région du monde hellénique à développer une civilisation florissante ; dès le VII[e] s. avant notre ère, l'ionien fut en Asie Mineure une langue littéraire. L'attique, dialecte de la ville d'Athènes, a produit aux V[e] et IV[e] s. une littérature merveilleuse, dont l'éclat n'a cessé de rayonner sur tout le monde civilisé.

b) le groupe *achéen*, dont il ne reste plus à date historique que trois tronçons épars dans l'arcadien, le cypriote et le pamphylien ; il n'est guère connu que par des inscriptions.

c) le groupe *éolien* ou du Nord-Est. Il comprend trois principaux dialectes, le thessalien, le béotien et le lesbien ; ce dernier a joué au VII[e]-VI[e] s. av. J.-C. un rôle important comme langue littéraire.

d) le groupe *dorien* ou occidental, qui comprend un grand nombre de parlers locaux assez différents les uns des autres et dont l'extension coïncide avec la grande invasion dorienne qui se répandit sur la Grèce en submergeant, semble-t-il, des populations achéennes. Parmi les parlers doriens figurent notamment ceux de Corinthe et de Mégare, de Laconie (avec les colonies de Tarente et d'Héraclée dans la Grande-Grèce), de Messénie, de Crète, de Sicile (Syracuse). Ce n'est guère qu'en Italie et en Sicile que s'est développée une littérature en dorien pur. La langue de la lyrique chorale, dont Pindare est le principal représentant, a un fond dorien, mais elle est composite.

Du groupe dorien se rapprochent quelques parlers du Nord-Ouest de la Grèce : Phocide (Delphes), Locride, Acarnanie, ainsi que le parler de l'Élide (Olympie) dans le Péloponnèse. Aucun n'a fourni de littérature.

1. A. Thumb, *Handbuch der griechischen Dialekte*, Heidelberg, 1909 ; C. D. Buck, *Introduction to the Study of the Greek Dialects*, Chicago, 1910.

Les dialectes si variés du grec ancien n'ont pas survécu. A partir du IVe s. avant notre ère, ils se sont tous peu à peu fondus dans la langue commune hellénistique (*hè koinè dialektos*), dont le dialecte attique constitue le fond [1]. On ne trouve donc pas trace des divisions dialectales anciennes dans le grec moderne. Celui-ci est sorti de la *koinè*, et s'il présente aujourd'hui une assez grande variété de parlers locaux, c'est que l'unité de la *koinè* s'est brisée à son tour. Les particularités des parlers actuels qu'on croit pouvoir faire remonter aux anciens dialectes sont extrêmement rares et de peu de portée [2].

Le plus ancien document du grec moderne peut être cherché dans le Nouveau Testament, dont la langue représente exactement le grec commun en usage dans le peuple. Mais il est malaisé de suivre jusqu'à nos jours le développement de cette langue parlée [3]. Le grec byzantin, langue officielle de l'empire d'Orient dont le siège était à Byzance (Constantinople) est une imitation artificielle du grec classique ; et de nos jours encore, depuis la libération de la Grèce, l'école, soutenue par l'église, s'efforce de maintenir comme langue écrite une langue « puriste » (*kaþarévousa*) rapprochée autant que possible de la *koinè* antique et très différente du grec parlé [4]. Celui-ci, qui porte en général le nom de *romaïque* ou roméique [5] présente des variétés dialectales qui n'ont été que partiellement étudiées [6]. Aujourd'hui le

1. A. Thumb, *Die griechische Sprache im Zeitalter des Hellenismus*, Strassburg, 1901 ; K. Dieterich, *Untersuchungen zur Geschichte der griechischen Sprache von der hellenistischen Zeit bis zum 10. Jahrhundert nach Christus*, Leipzig, 1898.

2. Elles sont à peu près limitées au tsakonien, dialecte du Péloponnèse parlé sur la côte orientale, à une quarantaine de kilomètres au Sud de Nauplie.

3. J. Psichari, *Essais de grammaire historique néo-grecque*, 2 vol. Paris, 1886-1887 ; du même, *Etudes de philologie néo-grecque*, Paris, 1892 ; G. N. Hatzidakis, *Einleitung in die neugriechische Grammatik*, Leipzig, 1892 ; Jannaris, *an Historical Greek Grammar chiefly of the Attic dialect as written and spoken from classical antiquity down to the present time*, London, 1897 ; H. Pernot, *D'Homère à nos jours*, Paris, 1921.

4. Voir K. Krumbacher, *Das Problem der neugriechischen Schriftsprache*, München, 1903.

5. L. Roussel, *Grammaire descriptive du roméique littéraire*, Paris, 1922.

6. P. Kretschmer, *Der heutige lesbische Dialekt verglichen mit den übrigen*

domaine du grec comprend outre la Grèce proprement dite, les îles de la mer Ionienne et celles de la mer Égée, de nombreux établissements en Asie Mineure [1] sur la côte et autour de Smyrne et même dans l'intérieur, jusqu'en Cappadoce (Césarée) ; sur la côte de la mer Noire, les colonies grecques s'étendent entre Samsoun et Trébizonde, et même au delà de cette ville. Il y a d'autre part d'anciens établissements grecs dans l'Italie du Sud (Terre d'Otrante) [2] et en Corse (Cargese), qui ont conservé l'usage de leur langue. Enfin, les Grecs ont fondé des colonies prospères sur de nombreux points du monde méditerranéen (Égypte) et en Amérique du Nord.

*
* *

Au Nord-Ouest de la péninsule des Balkans et faisant en quelque sorte la liaison par le continent entre le monde grec et le monde italique se trouvait placé dans l'antiquité le groupe ILLYRIEN. Il est des plus mal connus. Quelques inscriptions mal déchiffrées sont tout ce qui en reste. Outre l'illyrien proprement dit, il comprenait, semble t-il, le *vénète* (voir p. 54) et dans l'Italie du Sud le *messapien* [3].

Il n'est pas sûr que l'ALBANAIS moderne représente l'ancien

nordgriechischen Mundarten, Wien, 1905 ; H. Pernot, *Études de linguistique néo-hellénique, I, Phonétique des parlers de Chio*, Paris, 1907 ; K. Dieterich, *Sprache und Volksüberlieferung der südlichen Sporaden im Vergleich der übrigen Inseln des ägäischen Meeres*, Wien, 1908.

1. R. M. Dawkins, *Modern Greek in Asia Minor*, Cambridge, 1916.

2. G. Morosi, *Studi sui dialetti greci della Terra d'Otranto*, Lecce, 1870 ; *Dialetti romaici del Mandamento di Bova in Calabria*, in *Archivio Glottologico*, vol. IV (1874) ; A. Pellegrini, *Il dialetto greco-calabro di Bova*, vol. I, Rome 1880 ; Tozer, *The Greek speaking population of Southern Italy*, dans le *Journal of Hellenic Studies*, t. X (1890) ; de Groutars, *Musée belge*, t. I et II.

3. Kretschmer, *Einleitung*, p. 244 et suiv. ; H. Pedersen, *Zeitschrift f. vergl. Sprachforschung*, XXXVI, 299 et suiv. ; Hirt, *Die Indogermanen*, II, 604. — Sur le vénète : Pauli, *Altitalische Forschungen*, t. I et III. Sur le messapien : Deecke, *Rheinisches Museum*, t. XXXVI, p. 576, XXXVII, p. 373, XL, p. 133, XLII, p. 226 ; A. Torp, *Indogerm. Forschungen*, V, p. 195.

illyrien [1] ou même qu'il descende d'une langue appartenant à ce groupe ; mais sa position géographique oblige à le ranger ici.

Parmi les dialectes indo-européens, l'albanais est des plus tardivement connus ; si l'on met à part les mots recueillis par Arnold von Harff (fin du XVe s.), les plus anciens textes n'en remontent guère plus haut que le XVIIe siècle [2].

Il y eut un dictionnaire latin-albanais imprimé à Rome en 1635, puis à la fin du siècle diverses publications d'ouvrages de piété et de textes religieux traduits en albanais. Au XIXe s., on a recueilli en Albanie un grand nombre de contes et de chansons populaires. La langue comprend aujourd'hui deux dialectes principaux, dont la limite est marquée par le cours du fleuve Chkoumbi : au Nord, le *guègue* parlé notamment par les Malissores et par les Mirdites ; au Sud, le *tosque*. Un parler albanais de type tosque est encore aujourd'hui en usage dans quelques villages de la Calabre [3]. En outre, il y a eu en Grèce des établissements albanais qui ont laissé des traces notamment en Attique (à Éleusis). Les États-Unis d'Amérique renferment enfin pas mal d'Albanais. On compte en tout 1.800.000 Albanais au maximum.

*
* *

Groupe italo-celtique.

A l'extrémité occidentale du domaine indo-européen se trouve placé le groupe ITALO-CELTIQUE. L'unité de ce groupe linguistique semble assurée par le nombre de particularités communes que présentent l'italique et le celtique en opposition avec les autres dia-

1. C'est pourtant la doctrine soutenue par Gustav Meyer, *Beitr. zur Kunde der indogerm. Sprachen*, t. VIII, p. 135 ; cf. Pedersen, *Zeitsch. f. vergl. Sprachforschung*, t. XXXVI, p. 277.

2. G. Meyer, *Albanesische Studien*, 6, parues dans les *Sitzungsberichte* de l'Académie de Vienne, 1883-1897 ; Meyer Lübke, *Grundr. der rom. Phil.*, t. I, p. 1038 ; A. Dozon, *Manuel de la langue chkipe ou albanaise*, Paris, 1878 ; Pekmezi, *Grammatik der alban. Spr.*, Wien, 1908 ; Weigand, *Alban. Grammatik*, Leipzig ; A. Straticò, *Letteratura albanese*, Milano, 1896 (manuali Hoepli).

3. Lucien Bonaparte, *Transactions of the Philological Association*, London, 1882-1884, p. 492 et ss.

lectes indo-européens [1]. Mais l'italique et le celtique n'apparaissent dans l'histoire que déjà fortement différenciés. On ne sait rien sur la façon dont les peuples qui parlaient ces langues sont venus s'installer dans les régions où on les rencontre à date historique. Il est probable que ces installations résultèrent de migrations successives échelonnées sur de vastes espaces de temps. Ceux que nous connaissons le mieux ne sont peut-être que les derniers venus ; et il a pu exister du groupe italo-celtique des représentants plus anciens, dont nous avons perdu la trace. Parmi ceux-ci figure peut-être le *ligure*, langue d'un peuple conquérant qui occupa jadis un vaste territoire en Italie et dans le Sud de la Gaule et qui fut peu à peu réduit au territoire de la Ligurie historique sur le golfe de Gênes [2]. Au ligure paraît se rattacher le *sicule*, parlé en Sicile antérieurement au grec et au latin [3]. On ne sait au juste s'il faut attribuer au ligure ou au vénète les inscriptions dites lépontiques, trouvées en Italie à l'E. du lac de Come [4] ; cela semble indiquer qu'il y avait des liens linguistiques assez étroits entre le ligure et le vénète, c'est-à-dire entre l'italo-celtique et le groupe illyrien.

L'ITALIQUE [5], vers l'an 400 av. J.-C., devait comprendre dans la péninsule trois langues principales : l'ombrien, le latin et l'osque [6].

A l'époque historique, l'*ombrien* apparaît comme la langue d'une étroite région, resserrée entre l'Apennin et la rive gauche du Tibre. Le domaine de cette langue devait être auparavant

1. A. Meillet, *Dialectes indo-européens*, p. 31 ; A. Walde, *Ueber älteste sprachliche Beziehungen zwischen Kelten und Italikern*, Innsbruck, 1917.

2. G. Dottin, *Anciens peuples de l'Europe*, Paris, 1916, p. 180 ; C. Jullian, *Histoire de la Gaule*, t. I, p. 110.

3. E. A. Freeman, *History of Sicily*, t. I, p. 125.

4. Pauli, *Altitalische Forschungen*, Leipzig, 1885 ; O. A. Danielsson, dans les *Skrifter utgifna af k. Humanistiska Vetenskaps-Samfundet i Uppsala*, t. XIII, 1 (1909) ; H. Pedersen, *Philologica*, I, p. 38.

5. H. Nissen, *Italische Landeskunde*, t. I, p. 502 et ss.

6. Sur les dialectes italiques autres que le latin v. R. von Planta, *Grammatik der oskisch-umbrischen Dialekte*, 2 vol. Strassburg, 1892-1897 ; C. D. Buck, *A grammar of Oscan and Umbrian*, Boston, 1904.

beaucoup plus vaste, notamment au Nord et à l'Ouest ; il s'est réduit sous la pression de langues voisines, en particulier de l'étrusque. Le principal et presque l'unique monument qu'on ait de l'ombrien consiste en sept tables de bronze sur lesquelles sont gravés des morceaux d'un rituel sacrificatoire. Elles ont été découvertes en 1444 à Gubbio (anc. Iguvium) et sont certainement antérieures à l'ère chrétienne ; elles sont connues sous le nom de tables Eugubines. On ignore à quelle date l'ombrien a cessé d'être parlé.

L'*osque* était la langue des Samnites, peuple montagnard, batailleur et turbulent, qui lutta longtemps contre Rome. Son domaine était assez vaste. Outre le Samnium, il embrassait le Bruttium, le Picenum, la Lucanie et enfin la Campanie, où il était sans doute encore parlé au Ier siècle de notre ère. On a trouvé à Pompéi et à Herculanum des inscriptions en langue osque; mais les principaux textes viennent de Capoue, d'Abella, de Bantia, d'Agnone et sont antérieurs à l'ère chrétienne. Quelques textes osques proviennent même de Sicile (Messine).

Un dialecte voisin de l'osque est le *pélignien* dont le centre était à Corfinium. Entre le domaine ombrien et le domaine osque s'échelonnaient d'ailleurs quelques dialectes, insuffisamment connus par de rares et courtes inscriptions, qui étaient parlés dans les régions montagneuses du centre de l'Italie : tels le *volsque*, le *marse*, le *marrucin*, le *sabellin*, etc. Ils ont été de bonne heure absorbés par le latin.

Le *latin* [1] est proprement le dialecte du Latium. Ce n'était même à l'origine que le dialecte de la ville de Rome, le « sermo urbanus » ; comme tel il s'opposait au « sermo rusticus » des territoires voisins, Lanuvium, Préneste ou Falerii [2]. Peu à peu il s'étendit à tout le pays alentour, gagna les régions montagneuses où se parlaient les menus dialectes ci-dessus mentionnés, conquit les domaines ombrien et osque, absorba même les langues non ita-

1. F. Stolz, *Geschichte der lateinischen Sprache*, Leipzig, 1910 ; le même, *Historische Grammatik der lateinischen Sprache*, Leipzig, t. I, p. 25 et ss.

2. A. Ernout, *Les éléments dialectaux du vocabulaire latin*, Paris, 1909.

liques de l'Italie, l'étrusque et le celtique au Nord, le messapien au Sud, et par une fortune vraiment inouïe finit par s'imposer à la plus grande partie du monde occidental, y compris la Gaule, l'Espagne, et l'Afrique du Nord [1]. Il est demeuré langue savante longtemps après qu'on eût cessé de le parler ; il sert encore de langue universelle à l'église catholique.

Le plus ancien texte latin est sans doute l'inscription que porte une fibule d'or trouvée à Préneste en 1871 ; on le fait remonter aux environs de l'an 600 avant J.-C. Très anciennes sont aussi l'inscription que porte un cippe mutilé trouvé dans le forum romain en 1899 et celle qui orne le vase dit de Duenos, trouvé en 1880 au pied du Quirinal. C'est surtout à partir du IIe s. que les inscriptions deviennent abondantes. A l'époque impériale elles sont innombrables, couvrant à peu près toute l'étendue du monde connu des anciens. Quant à la littérature latine, aussi bien classique que médiévale, païenne que chrétienne, elle ne le cède en importance qu'à la seule littérature grecque.

La colonisation étendit l'usage de la langue latine à toutes les parties de l'empire romain. Lorsque l'unité de l'empire se brisa, la langue commune qui était partout en usage se brisa aussi en un certain nombre de parlers locaux [2].

A l'époque latine succède alors l'époque romane. Parmi les PARLERS ROMANS, certains sont demeurés jusqu'à nos jours à l'état de patois ou ont été absorbés par des langues voisines. Mais quelques-uns ont tenté de s'élever à la dignité de langues littéraires et, parmi les dialectes romans, il y en a qui sont devenus des langues communes. La constitution d'États politiquement unifiés a favorisé la diffusion de grandes langues communes, qui souvent avaient déjà depuis longtemps une existence comme

1. A. Budinsky, *Die Ausbreitung des lateinischen über Italien und die Provinzen des römischen Reiches*, Berlin, 1881.

2. H. Schuchardt, *Vokalismus des Vulgärlateins*, Leipzig, 1866-1869 ; F. G. Mohl, *Introduction à la chronologie du latin vulgaire*, Paris, 1899 ; E. H. Grandgent, *An introduction to vulgar Latin*, Boston 1907 ; J. Pirson, *La langue des inscriptions latines de la Gaule*, Bruxelles, 1901 ; A. Carnoy, *Le latin d'Espagne d'après les inscriptions*, 2e éd., Bruxelles, 1906.

langues littéraires. Il y a donc entre les différentes langues romanes des différences de valeur littéraire ou culturelle, d'importance politique et d'extension [1]. On peut ramener les différents parlers romans aux groupes suivants [2] :

a) groupe italien [3], comprenant un grand nombre de dialectes variés depuis la Vénétie jusqu'à la Sicile, y compris la Sardaigne et la Corse ; l'italien commun est essentiellement une langue littéraire développée dans la société cultivée de la ville de Florence, lingua toscana. Le plus ancien document en italien est de 960.

b) groupe provençal [4], comprenant outre les parlers de la Provence ceux du Languedoc, de l'Auvergne, du Limousin, de la Gascogne. La littérature provençale commence au XIe siècle.

c) groupe français [5], où l'on distingue un français du Nord (poitevin, normand et anglo-normand, picard, wallon, lorrain,

1. Sur les langues romanes en général, voir W. Meyer-Lübke, *Grammatik der romanischen Sprachen*, Leipzig (trad. fr. par Rabiet et Doutrepont, 4 vol., Paris, 1890-1906). A consulter aussi E. Bourciez, *Éléments de linguistique romane*, Paris, Klincksieck, 1910 (2e édition sous presse), et G. Millardet, *Linguistique et dialectologie romanes*, Paris, Champion, 1923.

2. Meyer-Lübke, *Einführung in das Studium der romanischen Sprachwissenschaft*, Heidelberg, 1909. Sur le nombre d'individus appartenant à chaque groupe, voir les données de statistique reproduites par Gröber, *Grundriss der romanischen Philologie*, t. I, p. 543 ; sur les plus anciens documents de chaque groupe dialectal, id., *ibid.*, p. 553 et ss.

3. W. Meyer-Lübke, *Italienische Grammatik*. Sur les dialectes : G. Ascoli, *L'italia dialettale* dans l'*Arch. glott. Ital.*, t. VIII, p. 98 ; et les travaux de d'Ovidio, Guarnerio, Salvioni, Ceci, Morosi, de Gregorio, Parodi, S. Pieri, etc. dans le même périodique. Résumé dans Giulio Bertoni, *L'Italia dialettale*, Milan (Hoepli), 1916.

4. Suchier, dans le *Grundr. der rom. Phil.*, t. I, p. 758 ; J. Ronjat, *Essai de syntaxe des parlers provençaux modernes*, Paris, 1913.

5. Foncin, *la langue française dans le monde*, Paris, 1900 (publication de l'Alliance française) ; F. Brunot, dans l'*Histoire de la langue et de la littérature française*, de Petit de Julleville, t. VIII, p. 861 (avec des cartes reproduites d'un travail de Gallois dans les *Annales de géographie*, t. IX, p. 211) ; le même, *Histoire de la langue française*, Paris, 5 vol. (en cours de publication) ; sur l'extension du français, H. Suchier, dans le *Grundr. der rom. Phil.*, I, p. 723 ; sur les dialectes, id., *ibid.*, p. 766 ; Gilliéron et Edmont, *Atlas linguistique de la France*, et les thèses ou travaux de O. Bloch, Ch. Bruneau, A. Devaux, A. Dauzat, Ch. Guerlin de Guer, G. Millardet, P. Rousselot, A. L. Terracher, etc.

bourguignon, champenois, îledefrançais) et un français du Sud-Est ou franco-provençal (lyonnais, dauphinois, vaudois, neuchâtelois, valaisien, savoyard). Le français commun est parti de l'usage de la bourgeoisie parisienne; il a été fixé au début du XVIIe siècle. Le plus ancien document de la langue française est fourni par les serments de Strasbourg (842).

d) groupe espagnol [1], avec des subdivisions en Léon, en Castille, en Aragon et en Andalousie. L'espagnol commun a pour base un dialecte du Nord de la Castille; il est parlé aux îles Canaries. Les plus anciens textes espagnols sont de 1145 (charte d'Oviedo) et 1155 (Fuero de Avilès).

e) groupe catalan [2], qui outre la Catalogne comprend la province de Valence et les îles Baléares. Le plus ancien texte est de 1171; il provient du monastère de Roda.

f) groupe portugais [3] comprenant aussi les parlers de la Galice, comme ceux des Açores et de Madère. Une charte de 1192 est le plus ancien texte portugais.

g) groupe rhéto-roman ou ladin [4] comprenant certains parlers des Grisons, du Tyrol et du Frioul.

h) groupe dalmate [5], le moins important de tous, représenté

1. G. Baist, *Die spanische Sprache*, dans le *Gr. der rom. Phil.*, t. I, p. 878; R. Menéndez Pidal, *Manual de gramática histórica Española*, 4e éd., Madrid, 1918; Gonçálves Viana, *Les langues littéraires de l'Espagne et du Portugal* dans la *Revue hispanique*, t. I, p. 1-21 (1894).

2. A. Morel-Fatio et J. Saroïhandy, *Das Catalanische* dans le *Grundr. der rom. Phil.*, I, p. 841.

3. Jules Cornu, *Die portugiesische Sprache*, dans le *Gr. der rom. Phil.*, t. I, p. 916; Leite de Vasconcellos, *Estudos de philologia mirandesa*, 2 vol., Lisbonne, 1900-1901; le même, dans la *Rev. Lusit.*, t. II et IV; le même, *Esquisse d'une dialectologie portugaise*, Paris, 1901.

4. G. Ascoli, *Saggi ladini* dans l'*Arch. glott. Ital.*, t. I; Theodor Gartner, *Rätoromanische grammatik*, 1883, et *Die rätoromanischen Mundarten*, dans le *Grundriss der roman. Phil.*, t. I, p. 608; Böhmer, *Verzeichniss rätoromanischer Literatur*, dans les *Romanische Studien*, t. VI (complété périodiquement par le *Kritischer Jahresbericht* de Vollmöller); R. Brandstetter, *Rätoromanische Forschungen*, 1905.

5. M. G. Bartoli, *Das Dalmatische* (Altromanische Sprachreste von Veglia bis Ragusa und ihre Stellung in der apennino-balkanischen Romania), dans les

par des parlers jadis en usage sur la côte de l'Adriatique, de l'île de Veglia jusqu'à Raguse. Tous ces parlers sont morts aujourd'hui ; le ragusain depuis le XVe s. ; le vegliote ne s'est éteint qu'à la fin du XIXe siècle.

i) groupe roumain [1] avec quatre principaux dialectes : le valaque, le moldave, le roumain de Transylvanie et le roumain du Banat. Il y a une littérature roumaine depuis le XVIe siècle. Le roumain est parlé aussi en plusieurs régions de la péninsule des Balkans ; il y a un groupe dialectal roumain comprenant certaines parties de l'Albanie, de la Thessalie, surtout de la Macédoine autour de Monastir et la vallée de Meglena, au Nord-Ouest de Salonique ; il y a enfin des restes de parlers roumains en Istrie.

Certaines des langues romanes ont pris par la colonisation une extension mondiale. Ainsi l'espagnol est la langue de l'Amérique du Sud [2] (moins le Brésil), de l'Amérique centrale (y compris la plus grande partie des Antilles) et du Mexique ; il est parlé dans une partie de la Californie et du Texas ; l'émigration l'a transporté sur la côte septentrionale de l'Afrique au Maroc et en Algérie (surtout dans la province d'Oran) ; sur le judéo-espagnol, voir page 60. Le portugais est la langue du Brésil ; il est aussi employé dans les colonies portugaises de l'Afrique et des Indes. L'italien est représenté en Amérique par les importantes colonies

Schriften der Balkankommission, t. IV et V, Wien, 1906 ; A. Ive, *Il dialetto veglioto* dans l'*Arch. glott. Ital.*, t. IX.

1. H. Tiktin, *Die rumanische Sprache* dans *Grundr. der roman. Phil.*, I, p. 564 ; O. Densusianu, *Histoire de la langue roumaine*, t. I (seul paru), Paris, 1901 ; sur les dialectes roumains, voir les articles de G. Weigand, dans le *Jahresbericht* de Vollmöller et son *Linguistischer Atlas des dakorumänischen Sprachgebietes*. Sur le roumain du Banat, Weigand, *Der Banater Dialekt*, 1896 ; sur le roumain de Macédoine, L. Lamouche, *Les Roumains de Macédoine*, Paris, 1913, G. Weigand *die Aromunen*, 1895, et Tache Papahagi, *Antologie Aromânească*, Bucarest, 1922 ; sur le roumain mégléníte, G. Weigand, *Vlacho-meglen*, 1892 ; sur le roumain d'Istrie, G. Weigand, *Romania*, t. XXI, p. 240 (1892).

2. R. Lenz, *Diccionario etimologico de las voces chilenas derivadas de lenguas indigenas americanas*, 1904-1905 ; R. J. Cuervo, *Apuntaciones criticas sobre el lenguaje Bogotano*, Paris, 5e éd., 1907.

établies en Argentine et aux États-Unis, en Afrique par celles de Tunisie, sans parler des possessions italiennes. On parle français dans une partie du Canada [1], dans quelques Antilles, dans une partie notable de l'Afrique du Nord, dans les colonies françaises d'Afrique (y compris Maurice), d'Océanie et d'Extrême-Orient. Le français est parlé par les gens cultivés de nombreux pays dans l'Europe centrale et sud-orientale, en Égypte et en Asie Mineure.

Sur plusieurs points du globe, le portugais, l'espagnol et le français, se trouvant en contact avec des langues indigènes, ont produit des parlers créoles, sortes de langues mixtes dont plusieurs ont été assez bien étudiées [2].

On peut évaluer à 53 millions le nombre des individus parlant espagnol, contre 48 millions qui parlent français et 39 qui parlent italien.

De même que les juifs d'Allemagne avaient un parler allemand (v. p. 69), de même les juifs d'Espagne (Sefardim) [3] en avaient un espagnol qu'ils emportèrent avec eux quand ils furent expulsés de la péninsule ; ce parler vit encore dans les communautés

1. Napoléon Legendre, *La langue française au Canada*, 1891 ; L. de Montigny, *La langue française au Canada*, Ottawa, 1916.

2. Addison van Name, *Contributions to creole grammar* (transactions of the American Philological Association, 1869-1870) ; F. A. Çoelho, *Dialectos romanicos ou neolatinos na Africa, Asia, e America* (*Boletin da Sociedade de geographia de Lisboa*, II, 3, p. 4-24) ; H. Schuchardt, *Sur le patois créole de la Réunion* (*Romania*, t. XI, p. 589), *Kreolische Studien*, dans les *Sitzungsberichte* de l'Académie de Vienne, 1882-1890, et dans la *Zeitschrift für romanische Philologie*, t. XII et XIII ; S. R. Dalgado, *Dialecto indoportugues de Ceylao*, 1900 ; A. Dietrich, *Les parlers créoles des Mascareignes* (Romania, t. XX, p. 216) ; C. Baissac, *Etude sur le patois créole mauricien*, Nancy, 1891.

3. Sur le judéo-espagnol, outre l'ouvrage général de H. Lœwe, *Die Sprachen der Juden*, voir : Grünwald, *Zur romanischen Dialectologie*, Heft 1, Belovar, 1883 ; Subak, *Zum Judenspanischen* (*Zeitsch. f. roman. philol.*, 1906). Sur la langue des Juifs, à Bordeaux : G. Cirot, *Bullet. hispan.* [1906], t. VIII, p. 172-189 ; au Maroc, Meneu, *Illustr. españ. y amer.*, 30 juillet 1890 ; Nahon, *Les Israélites du Maroc* (*Revue des études ethnogr. et sociolog.*, sept-oct. 1909) ; à Constantinople, Wagner, *Beiträge zur Kenntniss des Judenspanischen von Konstantinopel* (*Sitzber. der Wiener Akademie*, 1914) ; à Salonique, L. Lamouche, *Quelques mots sur le dialecte espagnol parlé par les Israélites de Salonique* (*Roman. Forschungen*, t. XXIII [Mélanges Chabaneau]).

juives de Salonique et de la Macédoine, de l'empire ottoman et de l'Afrique du Nord (côte nord du Maroc). Le judéo-espagnol et le judéo-portugais furent même parlés en Hollande et plus au Nord encore; il y en a quelques traces à Amsterdam et à Hambourg [1].

Enfin il faut mentionner les langues spéciales en usage dans les ports de la Méditerranée et qu'on désigne du nom de *sabir* ou de *lingua franca* [2] : l'élément roman (français, italien, espagnol) y tient une place prépondérante.

*
* *

Le CELTIQUE [3] a eu dans l'histoire des destinées moins brillantes que l'italique.

Il y eut sans doute une langue commune parlée par les populations celtiques qui, dans les siècles précédant l'ère chrétienne, ont occupé le centre de l'Europe, la Gaule, l'Italie du Nord, l'Espagne et qui ont poussé jusqu'en Grèce et en Orient, où elles ont fondé le royaume de Galatie. Mais de cette langue nous ne connaissons guère, en dehors d'un grand nombre de noms propres conservés dans la littérature classique ou dans des documents épigraphiques, qu'une soixantaine d'inscriptions, généralement fort courtes, trouvées dans le Nord de l'Italie, dans la vallée du Rhône et ailleurs encore en France. Les unes sont en caractères étrusques ; ce sont de beaucoup les moins claires. Les autres sont en caractères grecs ou latins. Les faits de langue qu'elles présentent sont trop peu importants pour permettre d'établir des distinctions dialectales.

1. Leite de Vasconcellos, *Esquisse d'une dialectologie portugaise*, p. 195.

2. H. Schuchardt, *Die Lingua Franca* (*Zeitsch. für roman. Philologie*, [1909], t. XXXIII, p. 441-461).

3. Voir en général E. Windisch, article *Keltische Sprachen* dans *l'Allgemeine Encyclopädie der Wissenschaften und Künste* de Ersch und Gruber, et *Grundriss der romanischen Philologie*, I, 283-312 ; E. C. Quiggin, article *Celt* dans *The Encyclopædia Britannica*, 11e éd. (Cambridge, 1910) ; H. Zimmer, *Die keltischen Sprachen* dans *die Kultur der Gegenwart*, I, XI (1909) ; G. Dottin, *Manuel pour servir à l'étude de l'antiquité celtique*, 2e éd., Paris, 1915. Sur le mouvement celtique contemporain, consulter H. Zimmer, *Preussische Jahrbücher*, vol. 92 (1898), p. 426, vol. 93 (1898), p. 59 et 294, vol. 99 (1900), p. 454.

Voilà en quoi se résume notre connaissance du VIEUX-CELTIQUE, qu'on appelle aussi *celtique continental* ou simplement *gaulois* [1]. Ce gaulois n'a laissé aucune descendance ; il s'est éteint en Gaule même dans les premiers siècles de l'ère chrétienne, étouffé par le latin.

Les seuls dialectes celtiques qui aient survécu apparaissent dans l'histoire à une date bien plus tardive. Ils appartiennent au celtique dit « insulaire » et se répartissent en deux groupes : le gaélique et le brittonique.

Le GAÉLIQUE, qui a son berceau en Irlande, est d'abord connu par des inscriptions en caractères « ogamiques » trouvées en Irlande et en Galles et qui peuvent remonter au v[e] s. de notre ère. Du VIII[e] s. au X[e] s., on a en gaélique d'Irlande ou *irlandais* de courts textes de caractère religieux ou des gloses ajoutées à des textes latins. A partir du XI[e] s. la littérature irlandaise est fort abondante et variée ; c'est une des plus riches de l'Europe médiévale. Beaucoup des morceaux qu'elle contient, en poésie comme en prose, sont d'une composition bien antérieure à l'époque des manuscrits où on les trouve.

A la suite de péripéties qu'expliquent les circonstances historiques, l'irlandais depuis le XVI[e] s. cède aux empiétements de l'anglais. Au cours du XIX[e] s. notamment, il a rapidement perdu du terrain. Il est aujourd'hui relégué dans un certain nombre d'îlots des régions rurales et montagneuses de l'île. Les parlers actuels ont entre eux de sérieuses différences ; on peut les ramener à trois groupes : l'un méridional, dans le Munster [2] (Comtés de Waterford, de Cork et de Kerry), l'autre occidental dans le Connaught [3] (Comtés de Galway et de Mayo, avec les îles d'Arran), le dernier septentrional dans le Donegal [4]. En 1911, il n'y avait

1. G. Dottin, *La langue gauloise*, Paris, 1920.

2. R. Henebry, *Contribution to the Phonology of Desi Irish* (Greifswald, 1901) ; J. Loth, *Rev. de Phonétique*, III, p. 317.

3. F. N. Finck, *Die Araner Mundart* (Marburg, 1899) ; G. Dottin, *Revue celtique*, t. XIV, p. 97.

4. E. C. Quiggin, *A Dialect of Donegal* (Cambridge, 1906); A. Sommerfelt, *The Dialect of Torr Co. Donegal*, Christiania, 1922.

plus en Irlande qu'environ 580.000 individus qui parlassent l'irlandais (dont 20.000 ne parlaient que l'irlandais). De sérieux efforts sont entrepris depuis une trentaine d'années pour faire revivre la vieille langue nationale.

Dès le v[e] s. de notre ère, les Irlandais avaient pénétré au Nord-Ouest de l'Angleterre et en Écosse (Scotia Minor). Dans ce dernier pays, ils introduisirent leur langue qui, à partir du XVI[e] s., surtout grâce à la Réforme, s'y développa en langue littéraire indépendante. Le domaine du *gaélique d'Écosse* [1] qui s'étendait jadis à toute la partie septentrionale du pays au N. de la Clyde est aujourd'hui restreint à quelques régions des Highlands. Il n'y avait plus en 1901 que 230.000 individus qui parlassent gaélique en Écosse (dont 28.000 ne parlaient que le gaélique) ; ce chiffre a dû plutôt diminuer depuis.

Dans l'île de Man, il y a encore aujourd'hui quelques centaines de personnes qui parlent un dialecte gaélique, le *mannois* ou *manx* [2], plus voisin d'ailleurs du gaélique d'Écosse que de l'irlandais.

Les circonstances politiques et sociales ont contraint un nombre considérable d'Irlandais à émigrer, surtout dans la seconde moité du XIX[e] s. Il y a des colonies irlandaises très importantes aux États-Unis d'Amérique et en Australie ; elles ont en général abandonné l'usage de leur langue nationale.

Le BRITTONIQUE était parlé en Grande-Bretagne avant l'invasion des Romains ; il y subsista après la conquête, mais l'occupation romaine, qui dura près de quatre siècles (de 43 à 410 de notre ère), agit fortement sur lui ; les plus anciens textes brittoniques présentent un état de langue déjà très évolué. A partir du v[e] s., les invasions germaniques refoulèrent peu à peu le brittonique vers les régions occidentales de la Grande-Bretagne (Pays

1. James Murray, *Revue celtique*, t. II, p. 180 (avec une carte) ; C. M. Robertson, *The Celtic Review*, vol. III à V (1906-1908) ; Henderson, *Zeitschrift für celtische Philologie*, vol. IV-V (1903-1905).

2. H. Jenner, *The Manx language* dans les *Transactions of the London Philological Society*, 1875.

de Galles et Cornwall) et le contraignirent même à passer la mer pour aller s'établir dans la « petite Bretagne » ou Bretagne armoricaine. De là, dès les plus anciens textes, une division tripartite du brittonique.

a) le *gallois* [1], parlé aujourd'hui encore au Pays de Galles par près d'un million d'individus. Attesté par des gloses à partir du VIII[e] s. de notre ère, il est connu au moyen âge par une littérature assez riche, originale surtout en poésie, et qui se continue d'une façon ininterrompue jusqu'à nos jours. Il y a aujourd'hui un gallois littéraire, dont la fixation remonte à la traduction de la Bible (par le Dr. Morgan en 1588), et de nombreux parlers locaux, qui se laissent ramener à deux groupes principaux : l'un septentrional (Anglesey, Carnarvon, Merioneth), l'autre méridional (Cardigan, Carmarthen, Clamorgan). Les Gallois, installés aux États-Unis d'Amérique, en Patagonie (vallée du Chubut) et dans divers Dominions conservent en général l'usage de la langue.

b) le *cornique* [2], parlé jadis en Cornwall et dont le domaine rejoignait primitivement le domaine gallois, a disparu à la fin du XVIII[e] s., succombant aux empiétements de l'anglais ; on en possède quelques gloses qui remontent au IX[e] s., un vocabulaire daté du XII[e] s. et, aux XVI[e]-XVII[e] s., une série de drames religieux.

c) le *breton* [3], introduit en Armorique aux V[e] et VI[e] s. de notre

1. John E. Southall, *Wales and her language*, London, D. Nutt, 1893 ; id., *The Welsh language Census of 1901*, Newport, 1904 (with a coloured linguistic map of Wales) ; Sir John Rhys, *The Welsh People*, London, Fisher Unwin, 4[e] éd., 1906, chap. XII (Language and Literature) ; sur l'histoire du gallois, J. Morris Jones, *A Welsh grammar historical and comparative*, part I (Oxford, 1913), avec les additions et corrections de J. Loth, *Rev. Celt.*, t. XXXVI et XXXVII. Sur les dialectes, diverses notes ont paru dans les *Transactions of the Guild of Graduates of the University of Wales* (Cardiff) ; voir aussi O. H. Fynes-Clinton, *The Welsh Vocabulary of the Bangor-district* (Oxford, 1913), et Meredith Morris, *a Glossary of the Demetian dialect* (Tonypandy, 1910).

2. V. les nombreux articles de J. Loth, *Revue celtique*, t. XVII à XXIV et XXXII à XXXVII ; H. Jenner, *Handbook of the Cornish language* (London, 1904) ; Jago, *Ancient language and dialect of Cornwall* (London, 1882).

3. P. Sébillot, *Revue d'ethnographie*, janvier 1886 ; J. Loth, *Revue celtique*,

ère, s'y est maintenu jusqu'à nos jours dans la partie occidentale de la péninsule, limitée par une ligne à peu près droite allant de Plouha à l'embouchure de la Vilaine. On y distingue quatre principaux dialectes, le trégorois, le léonard, le cornouaillais et le vannetais, ce dernier assez différent des trois autres ; chacun présente d'ailleurs de nombreuses variétés. En dehors de gloses anciennes et de chansons populaires modernes, on a en breton des mystères et des ouvrages d'édification, qui commencent au XVIe s. et dont la tradition se poursuit jusqu'à nos jours. Il y a aujourd'hui encore environ 1.300.000 bretonnants ; mais le français est compris sur la plus grande partie du domaine breton.

*
* *

Groupe germanique.

Les écrivains de l'antiquité classique désignent sous le nom de Germains des peuples qui occupaient les plaines de l'Europe septentrionale entre la Vistule, le Rhin et les Alpes. C'est d'abord par l'intermédiaire des Celtes qu'ils en avaient eu connaissance ; le nom même de Germain semble avoir été celui d'une tribu germanique avec laquelle les Celtes se trouvaient en contact sur les bords du Rhin [1]. Au point de vue linguistique, le germanique [2] comprend trois groupes distincts :

a) le germanique oriental, représenté par le *gotique* [3]. Les

t. XXIV, p. 295, et t. XXVIII, p. 374 ; F. Vallée, *La langue bretonne en quarante leçons* (St-Brieuc, 4e éd., 1916) ; sur l'histoire du breton, consulter surtout J. Loth, *Chrestomathie bretonne*, Paris, 1890 ; E. Ernault, *Glossaire moyen-breton*, Paris, 1895-1896.

1. J. Carcopino, *Revue Celtique*, t. XXXVIII, p. 319.

2. H. Hirt, *Die Stellung des germanischen im Kreise der verwandten Sprachen Zeitschrift für deutsche Philologie*, t. XXIX, p. 289 et ss. ; F. Kluge, *Die sprachliche Stellung der Germanen, Internationale Wochenschrift*, t. V (1911), p. 721 et ss. ; le même, *Urgermanisch, Vorgeschichte der altgermanischen Dialekte*, 3e éd., Strassburg, 1913 ; A. Meillet, *Caractères généraux des langues germaniques*, Paris, 2e éd., 1923.

3. W. Streitberg, *Gotisches Elementarbuch*, 5e éd., Heidelberg, 1920.

Gots, qui occupèrent à date préhistorique la partie Sud de la Scandinavie, où le nom de Gothie se conserve encore, descendirent du bassin de la Vistule vers le Sud-Est en direction de la mer Noire. Au début du IIIe s. de notre ère, ils fondèrent entre le Bas-Danube et le Dniepr un royaume bientôt scindé en deux, celui des Ostrogots et celui des Wisigots. Ces deux peuples conquérants firent des expéditions vers l'Occident. Il y eut un royaume ostrogot en Italie jusqu'en 555 et un royaume wisigot en Espagne jusqu'en 711; mais leur langue y fut rapidement noyée dans le milieu roman qui l'entourait. En Orient, le gotique fut plus vivace, bien qu'il ait complètement disparu depuis environ trois siècles. En 1560 le Hollandais Angerius von Busbeck recueillit en Crimée des restes de langage gotique. En dehors de ce gotique de Crimée et de quelques textes fort courts découverts en Italie le gotique n'est connu que par les fragments d'une traduction de la Bible que fit pour l'édification de ses ouailles l'évêque Ulfila (né vers 311, mort en 383); son diocèse comprenait la Mésie inférieure, au pied de l'Hémus, avec la ville de Nicopolis. Ulfila écrivit sa traduction dans un alphabet emprunté principalement au grec, mais auquel se mêlent quelques caractères latins et même de vieux caractères runiques. — Au germanique oriental devaient appartenir aussi la langue des Burgondes et celle des Vandales; nous n'en avons guère conservé que quelques noms propres.

b) le germanique septentrional ou *nordique*, parlé en Scandinavie. C'est la langue germanique dont on possède les plus anciens documents. Le nordique est en effet connu dès le IIIe s. de notre ère par une soixantaine d'inscriptions écrites principalement en un alphabet spécial, dit runique, qui fut également employé par les Gots et par les Anglo-Saxons [1]. On désigne parfois la langue de ces inscriptions du nom de *vieux-norrois*. Elles s'échelonnent sur un espace d'environ cinq siècles. Mais il

1. Sophus Bugge, *Norges Indskrifter med de oeldre Runer*, Christiania, 1891; L. Wimmer, *Die Runenschrift*, Berlin, 1887; Maurice Cahen, *M. S. L.*, XXIII, p. 1 et ss.

faut ensuite descendre bien plus bas dans l'histoire pour rencontrer de nouveaux textes en nordique. La langue conserva sans doute une certaine unité jusqu'au VIIIe s. ; du VIIIe au XIe s. s'étend une période que marque l'extension des peuples scandinaves dans trois directions et en même temps la scission du nordique en trois groupes différents. Les Norvégiens s'étendent jusqu'en Écosse et en Irlande ; ils prennent les Shetlands, les Orcades, les Hébrides, les Féroé, et font vers l'an 900 la conquête de l'Islande. Les Danois descendent en Slesvig, passent en Angleterre où ils s'établissent surtout à l'Est et au Nord, et vont jusqu'en Irlande ; en Normandie, le danois était encore parlé à Bayeux au XIIe siècle. Les Suédois s'étendent en Finlande, en Estonie, en Livonie, en Russie même où ils occupèrent Novgorod jusque vers 1300. Le nordique comprenait dès le XIe s. quatre groupes de langues [1] : le norvégien, le danois, le suédois et l'islandais, celui-ci n'ayant pas tardé à évoluer indépendamment du norvégien.

C'est sous la forme de l'*islandais* que le nordique devait au moyen âge briller du plus vif éclat. Dès le XIIe s., il y eut en Islande une activité littéraire intense et une production d'une importance exceptionnelle ; des légendes (sagas) furent réunies et mises en vers sous le nom général d'Edda, des textes juridiques (grâgâs) furent rédigés, ainsi que des biographies d'hommes illustres. C'est surtout aux sagas que le vieil islandais doit sa célébrité. Il y a aujourd'hui encore une littérature islandaise assez active.

Le *norvégien* est connu à une date à peine moins ancienne que l'islandais ; en tant que langue littéraire, il meurt à la fin du moyen âge. Le *suédois* et le *danois*, dont les plus anciens textes littéraires remontent au XIIIe s., n'ont pas cessé de vivre brillamment jusqu'à nos jours. Un danois adapté sous le nom de *riks-*

1. A. Noreen, *Geschichte der nordischen Sprachen*, dans le *Grundriss der germanischen Philologie*, t. I, p. 417 et ss. ; Petersen, *Det danske, norske og svenske Sprogs Historie*, Copenhague, 1829-1830 ; Dahlerup, *Det danske Sprogs Historie*, 2e éd., Copenhague, 1921 ; Marius Kristensen, *Nydansk*, ibid., 1906 ; Axel Kock, *Undersökningar i svensk sprakhistoria*, Lund, 1887 (Arkiv f. nord. Fil., IV, 163 et ss.).

maal a servi au XIXe s. de langue littéraire à la Norvège; des efforts sont faits actuellement pour lui substituer une langue nationale (*landsmaal*), fondée sur des parlers norvégiens [1].

c) le germanique occidental ou *westique*. Le westique est connu à une époque où il avait perdu déjà toute unité. Les deux principaux représentants en sont l'allemand et l'anglais.

L'*allemand* [2] comprend lui-même le haut-allemand et le bas-allemand. Le haut-allemand, dont les plus anciens documents sont des gloses du VIIIe siècle, se divise déjà à cette époque en de nombreux parlers locaux qu'on peut ramener à trois groupes principaux, le bavarois, l'alaman ou alémanique (tous deux réunis souvent sous le nom de haut-allemand au sens étroit, strenghochdeutsch ou oberdeutsch) et le francique ou franconien. Le domaine bavarois comprend, outre la Bavière, l'Autriche, la Styrie, une partie du Tyrol et de la Carinthie; plus au Sud, il a un îlot à Gottschee, en domaine slovène. Le domaine alaman se divise en haut-alaman (Suisse alémanique), bas-alaman (Bade et Alsace) et souabe (Würtemberg). Le domaine francique est celui qui présente les plus grandes variétés; plusieurs parlers franciques, qui ont de nombreux traits communs avec le bas-allemand, sont parfois réunis sous le nom de moyen-allemand (mitteldeutsch). C'est principalement sur des parlers franciques qu'a été constitué l'allemand commun [3]. Dès la fin du moyen

1. Achille Burgun, *Le développement linguistique en Norvège depuis 1814*, 2 vol. Christiania, 1919-1921.

2. O. Behaghel, *Geschichte der deutschen Sprache*, Strassburg, 1911 (avec une très complète bibliographie); le même, *Die deutsche Sprache*, 5e éd., Wien et Leipzig, 1911; H. Lichtenberger, *Histoire de la langue allemande*. Paris, 1895; sur les dialectes de l'Allemagne actuelle, voir surtout O. Behaghel, *Schriftsprache und Mundart*, Giessen, 1896; F. Wrede, *Deutsche Dialektgeographie, Berichte und Studien über G. Wenkers Sprachatlas des deutschen Reiches*, Marburg, 1908 et ss.; O. Weise, *Unsere Mundarten, ihr Werden und ihr Wesen*, Leipzig et Berlin, 1910; P. Kretschmer, *Wortgeographie der hochdeutschen Umgangsprache*, 1918.

3. A. Socin, *Schriftsprache und Dialekte im deutschen nach Zeugnissen alter und neuer Zeit*, Heilbronn, 1888; F. Kluge, *Unser Deutsch*, Strassburg; le même, *Von Luther bis Lessing*, 3e éd., Strassburg 1897; O. Weise, *Unsere Muttersprache, ihr Werden und ihr Wesen*, 7e éd., Leipzig, 1909.

âge la chancellerie impériale tendait à se créer une langue commune qui fût comprise dans tous les pays de son ressort. Cette langue, sanctionnée dans les milieux protestants par les écrits de Luther, est devenue l'allemand littéraire ; elle s'est rapidement étendue aux régions protestantes de l'Allemagne du Nord, puis a été adoptée dans le Sud, resté catholique. Elle a été transportée par la colonisation dans plusieurs pays d'Europe [1] (Hongrie, Roumanie, Russie); elle est également employée dans les groupements allemands de l'Amérique du Nord (États-Unis) et du Sud (Brésil, Argentine, etc.) [2]. Il peut y avoir sur terre 80 millions d'individus qui se servent de l'allemand [3].

Aux dialectes du haut-allemand il faut joindre le *yidich* [4] (jüdisch-deutsch), langue spéciale qu'employèrent les juifs Aškenazim, surtout à partir du XIVe s. où ils furent persécutés en Allemagne et expulsés du pays. On peut distinguer deux groupes de dialectes yidich : 1° un groupe oriental comprenant aujourd'hui tous les parlers employés de la Baltique à la mer Noire par les communautés juives de Lituanie, de Pologne, de Russie et en partie de Roumanie ; ces parlers sont très proches les uns des autres et sortent évidemment d'un ancêtre commun. Ils ont été transportés par l'émigration aux États-Unis où le yidich oriental se parle et s'écrit abondamment ; la seule ville de New York compte plus d'un million de juifs ; 2° un groupe occidental, qui a dû avoir aussi dans le passé une vaste extension, mais qui est aujourd'hui à peu près restreint aux communautés juives d'Alsace ; en Lorraine, dans la campagne de langue française autour

1. H. Nabert, *Das deutsche Sprachgebiet in Europa und die deutsche Sprache sonst und jetzt*, Stuttgart, 1893.

2. Paul Langhans, *Deutscher Kolonialatlas*, Gotha, 1893-1897.

3. Le *Geographisch-statistischer Universal-atlas* de Hickmann, portait ce chiffre à 90 millions, en 1912.

4. Heinrich Lœwe, *Die Sprachen der Juden*, Köln, 1911 ; L. Sainéan, *M.S.L.*, t. XII, p. 90 et 176 ; Mathias Mieses, *die Entstehungsursache der jüdischen Dialekte*, Wien, 1915 ; Gerzon, *Die jüdisch-deutsche Sprache*, Köln, 1902 ; Ernest Lévy, *M. S. L.*, XVIII, 317. Une bibliographie des ouvrages relatifs au yidich par Alfred Landau se trouve dans la revue *Deutsche Mundarten* (herausgegeben von W. Nagl, Wien), t. I (1896), p. 126.

de Metz, il achève de mourir. Il est complètement mort en Allemagne depuis le XVIII[e] s. où sous l'influence de Mendelssohn les juifs se sont mis à parler et à écrire l'allemand commun, sans renoncer d'ailleurs à l'emploi des caractères hébraïques. On n'a pas réussi à localiser le dialecte allemand d'où est sorti le yidich, aussi bien oriental qu'occidental ; il est toutefois hors de doute que c'est un dialecte haut-allemand de la région centrale, c'est-à-dire un dialecte francique. Le yidich d'Alsace, qui se maintient dans un milieu où l'on parle alaman, présente certains traits communs avec ce dialecte.

Le *bas-allemand*, dont le plus ancien texte est le poème du Hêliand (« Sauveur »), composé vers 830, est aujourd'hui représenté par une foule de parlers locaux désignés sous le nom général de *plattdeutsch*. La limite dialectale du haut et du bas-allemand suit une ligne à peu près droite coupant l'Allemagne de l'Ouest à l'Est et passant par Aix-la-Chapelle, Cassel, Nordhausen, Wittenberg et Schwiebus. La langue du Hêliand est généralement appelée *vieux-saxon*, du nom du peuple saxon qui occupait alors la région comprise entre Rhin et Elbe.

Au bas-allemand se rattachent le *hollandais* et le *flamand* [1]. Les deux langues n'en font qu'une ; elles sont issues d'un mélange de dialectes bas-allemands, que des conquérants Francs et Saxons avaient apportés avec eux. Suivant les hasards de l'histoire, il y eut à plusieurs reprises des créations de langues écrites, plus ou moins vivaces, sur divers points du pays, jusqu'à ce que la traduction de la Bible, entreprise de 1619 à 1639, fixât une norme définitive. Toutefois, les divisions religieuses et politiques créèrent en ces derniers siècles entre le flamand et le hollandais des différences, surtout de vocabulaire, que certains s'efforcent de maintenir aujourd'hui. Le hollandais a été transporté par la colonisation dans les Indes orientales (Insulinde) et occidentales (Antilles) ainsi que dans l'Afrique du Sud (Transvaal, Orange et partie de la Colonie du Cap) [2].

1. Jan te Winkel, *Geschichte der niederländischen Sprache*, dans le *Grundriss der germanischen Philologie*, I, p. 634-649 (1891).

2. H. Meyer, *Die Sprache der Buren*, Göttingen, 1901.

Des parlers différents du bas-allemand s'emploient aujourd'hui encore dans la partie septentrionale de la Hollande et dans les îles du Zuyderzée, puis sur la côte de la mer du Nord en quelques points du Oldenbourg, dans l'île de Héligoland et sur la côte du Slesvig au Nord de Husum ainsi que dans les îles voisines, jusqu'à Sylt. Ce sont les restes du *frison* [1], dialecte westique apparenté au vieux-saxon ; le frison fut jadis dans les Pays-Bas le rival du francique et du saxon et il fut au nombre des dialectes germaniques que l'invasion porta en Grande-Bretagne ; le northumbrien présente avec lui des traits communs.

L'*anglais* [2] a, dès l'origine, encore moins d'unité que l'allemand. Au cours des ve et vie s. de notre ère diverses tribus du Nord de la Germanie appartenant notamment aux Angles, aux Saxons, aux Frisons et aux Jutes envahirent par émigrations successives les parties orientales et méridionales de la Grande-Bretagne, alors occupée par des Celtes (Bretons), et y fondèrent des établissements qui s'organisèrent plus tard en ce qu'on appela l'heptarchie anglo-saxonne. On désigne souvent leur langue sous le nom d'anglo-saxon, ou de vieil-anglais, mais cette langue n'est que le groupement de dialectes variés, dont les principaux sont le northumbrien au Nord et le mercien au Centre, le saxon au Sud et le kentien à l'extrémité sud-est. Les plus anciens documents de l'anglo-saxon sont des gloses, dont quelques-unes paraissent antérieures au viiie s. La poésie a surtout fleuri sur le domaine angle et la prose sur le domaine saxon. Le danois, puis surtout le français exercèrent sur l'anglais une forte et durable influence ; et dès la seconde moitié du xive siècle il se développa en Angleterre une langue commune, partie de la ville de Londres

1. Th. Siebs, *Geschichte der friesischen Sprache*, *Gr. germ. Phil.*, p. 723 et ss.

2. F. Kluge, *Geschichte der englischen Sprache*, ibid., p. 780 et ss.; R. Huchon, *Hist. de la langue angl.*, Paris, 1923 ; O. Jespersen, *Growth and Structure of the English language*, 3e éd., Leipzig, 1919 ; Morsbach, *Ueber den Ursprung der neuenglischen Schriftsprache*, Heilbronn, 1888 ; Paul Horn, *Historische neuenglische Grammatik*, Strassburg, 1908 ; O. F. Emerson, *History of the English Language*, New-York, 1894 ; Henry Cecil Wyld, *A short History of English*, London, 1921 ; sur les dialectes, v. A. J. Ellis, *On Early English Pronunciation*, 5 vol. 1869-1889, et J. Wright, *The English Dialect grammar*, Oxford, 1905.

et constituée surtout d'éléments dialectaux du centre du pays. Cette langue s'étend aujourd'hui à l'île entière (moins le Pays de Galles et un coin des Highlands, v. p. 63) ; c'est peu à peu qu'elle a absorbé le cornique, fait reculer le gaélique d'Écosse, conquis les Orcades et les Shetlands (où l'on a parlé scandinave jusqu'au XVIe s.) et gagné l'Irlande (où la lutte avec l'idiome gaélique présente des péripéties très variées qui ne sont pas encore terminées). Les dialectes anglais de l'Écosse ont des formes spéciales d'un type archaïque [1] ; et sur un point de l'Irlande, au Sud du Comté de Wexford (baronies de Forth et de Bargy), s'est conservé jusqu'au milieu du XIXe s. un dialecte anglais, introduit là au XIIe s. et resté très aberrant [2]. Mais en général sur les domaines celtiques où l'anglais s'étend, c'est sous la forme de langue commune ; cela n'exclut pas d'ailleurs certaines formations particulières [3].

Hors d'Europe, l'anglais commun a conquis un domaine immense, par suite de la colonisation. Il occupe la plus grande partie de l'Amérique du Nord, aux États-Unis et au Canada, et, en Océanie, l'Australie et la Nouvelle-Zélande ; dans l'Afrique du Sud, en Égypte et aux Indes, l'anglais est plus ou moins la langue officielle, du commerce et de l'administration. Il se comprend dans une partie considérable du monde ; on évalue en gros à 150 millions les individus qui parlent anglais sur la surface du globe. Cela ne va pas sans entraîner des altérations. Il y a des prononciations spéciales qui se fixent et s'étendent, si bien qu'on peut reconnaître à son langage un Canadien ou un Australien. Aux États-Unis, l'anglais prend de plus en plus

1. Murray, *The Dialect of the Southern Counties of Scotland*, 1873 ; W. Grant et J. M. Dixon, *Manuel of Modern Scots*, Cambridge, 1921 ; sir James Wilson, *Lowland Scotch*, Oxford, 1915.

2. Jacob Poole (mort en 1827), *A glossary, with some pieces of verse, of the Old Dialect of the English Colony in the baronies of Forth and Bargy* (édité par William Barnes, 1867).

3. P. Joyce, *English as we speak it in Ireland*, Dublin, 2e édition, 1910 ; J. H. Staples, *Notes on Ulster English Dialect*, dans les *Transact. of the Philological Society* (1895-1898), p. 357.

certains caractères particuliers [1]. Il y a enfin des combinaisons d'anglais et de langues indigènes, comme le pidgin english d'Extrême-Orient ou le broken-english de la Sierra Leone : ce sont des sortes de langues créoles [2].

*
* *

GROUPES BALTIQUE ET SLAVE.

Malgré les différences sérieuses qu'elles présentent, on a pris l'habitude de grouper ensemble les langues baltiques et les langues slaves, qui s'opposent en effet aux groupes voisins par un certain nombre de traits communs.

Les LANGUES BALTIQUES sont au nombre de trois : le vieux-prussien [3], le lituanien [4] et le lette [5].

Le *vieux-prussien* est mort au XVIIe s. ; il est connu seulement par un vocabulaire dit « vocabulaire d'Elbing » (800 mots), qui remonte au XVe s. et par une traduction faite en 1561 de trois petits catéchismes et de l'Enchiridion de Luther.

Le *lette* et le *lituanien* sont encore vivants aujourd'hui. Le premier est parlé par 1.300.000 individus (en majorité luthériens), le second par 2.500.000 (presque tous catholiques). Le domaine du lette est la Courlande et la partie sud de la Livonie. Celui du lituanien est l'ancienne Lituanie russe, mais il déborde

1. H. L. Mencken, *The American language*, 2e éd., New-York, 1921.

2. C. G. Leland, *Pidgin english*, 5e éd., 1900 ; sur le broken-english de la Sierra-Leone, F. W. H. Migeod, *The languages of West Africa*, London, 1911-1913.

3. Berneker, *Die preussische Sprache* (Strassburg, 1896), et R. Trautmann, *Die altpreussischen Sprachdenkmäler* (Göttingen, 1910).

4. F. Kurschat, *Grammatik der littauischen Sprache* (avec une carte du domaine), Halle, 1876 ; Bezzenberger, *Beiträge zur Geschichte der litauischen Sprache*, Göttingen, 1877 ; id., *Zur litauischen Dialektforschung*, dans les *Beiträge zur Kunde der indogerm. Spr.*, vol. VIII, IX, X et XX.

5. Bielenstein, *Die lettische Sprache* (Berlin, 1863) ; J. Endzelin, *Lettische Grammatik*, Riga, 1922.

au delà de l'ancienne frontière prussienne, englobant les régions de Memel et de Tilsitt.

Le lette et le lituanien sont connus depuis le XVI[e] s. Les premiers documents qu'on en possède sont des traductions du catéchisme de Luther ; la traduction lituanienne date de 1547, la traduction lette de 1586. Encore au XIX[e] s. c'étaient surtout des langues de paysans ; cependant les événements politiques des dernières années ont eu pour résultat de leur conférer la dignité de langues nationales ; aussi tendent-ils à se répandre dans les villes. Une notable partie de la population de Riga parle lette.

Les deux langues comportent des variétés dialectales. On donne en particulier le nom de žémaite au dialecte lituanien parlé au Nord du domaine, entre Kovno et Chavli. Il s'est créé au XIX[e] s. une langue littéraire lituanienne reposant surtout sur les dialectes du Sud et qui a été par l'émigration transportée aux États-Unis, où nombre de publications en lituanien ont vu le jour dans le dernier quart de siècle. Le lituanien est remarquable par son archaïsme ; bien que le système général de la langue se soit aussi complètement transformé que celui des autres groupes de la famille, on y retrouve conservés, à l'état de survivances, un grand nombre de traits anciens. Le lette est en général beaucoup plus évolué ; ce qui tient en partie à ce qu'il s'est étendu sur des régions où se parlaient des dialectes finnois, notamment le live.

Les LANGUES SLAVES [1] se répartissent en trois groupes, méridional, occidental et oriental.

A. Le *slave méridional* est le plus anciennement connu. Lorsqu'au IX[e] s. de notre ère les apôtres Cyrille et Méthode voulurent traduire les textes sacrés pour les besoins de l'évangélisation, à la demande d'un prince slave de Moravie, ils prirent comme base de leur traduction leur dialecte maternel, qui était celui de la région de Salonique. La langue en laquelle cette traduction a été

1. L. Niederle, *La race slave*, Paris, 1911 (traduit du tchèque par L. Leger) et *Manuel de l'antiquité slave*, I (Paris, 1923).

faite est ce qu'on appelle le *vieux-slave*, ou vieux-slavon, ou encore vieux-slave ecclésiastique ; elle est restée durant tout le moyen âge la langue religieuse des Slaves orthodoxes et par suite a été comme telle lue et écrite par beaucoup de Slaves (Bulgares, Serbes, Russes), dont le parler était sensiblement différent.

Aujourd'hui le slave méridional s'étend de l'Adriatique à la mer Noire sur un domaine où les parlers locaux se succèdent suivant une gradation tellement peu sensible qu'il serait souvent impossible de marquer des divisions dialectales si l'on s'en tenait à la langue parlée. Il s'est cependant constitué avec le temps trois langues littéraires qui répondent en gros à trois types dialectaux différents :

a) le *slovène* [1], employé par environ 1.200.000 individus, dont le plus ancien texte est du xᵉ s. (fragment de Freising) et qui a une littérature depuis la fin du xviiiᵉ s. Le centre du domaine slovène est la Carniole, mais le slovène s'étend à la Styrie et à la Carinthie méridionales, à un coin de la Croatie et à la partie principale de l'ancien Küstenland autrichien jusqu'à l'Isonzo et même au delà de ce fleuve.

b) le *serbo-croate* [2], servant à environ neuf millions d'individus, répartis entre la Serbie proprement dite, la Croatie, la Bosnie-Herzégovine, le Monténégro, la Dalmatie, la Voïvodine (Banat, Bačka, Baranja) ; presque tous sont aujourd'hui réunis dans le royaume des Serbes, Croates et Slovènes. D'après la manière dont s'exprime le mot « quoi ? » on y distingue trois dialectes : le *štokavien* (qui est le principal et a servi de base à la langue littéraire), le *tšakavien* et le *kaykavien*.

Il y eut jadis des colonies serbo-croates sur la côte de l'Adriatique en Italie. Dans la province de Campobasso (contado di Molise) trois villages ont encore aujourd'hui une population

1. Ivan Krek, *Les Slovènes*, Paris, Alcan, 1917 (traduit par A. U.).

2. Voir Olaf Broch, *Die Dialekte des südlichen Serbiens*, et M. Rešetar, *Der štokavische Dialekt*, dans les *Schriften der Balkan-Commission, Linguistische Abteilung*, Vienne, 1905 et 1907 ; A. Leskien, *Grammatik der serbo-kroatischen Sprache*, Heidelberg, 1914 ; A. Belić, *Zum heutigen Stande der serbo-kroatischen Dialektologie*, dans le *Rocznik Sławistyczny* (1910), III, 82.

qui parle un dialecte štokavien [1]. Le serbo-croate s'écrit avec deux alphabets : le cyrillique chez les orthodoxes, le latin chez les catholiques. Bien qu'il ait eu déjà à partir du XVIe s., et notamment au XVIIe s., une littérature en Dalmatie et à Raguse, c'est surtout au XIXe s. que le serbo-croate a pris la valeur d'une langue littéraire et d'une langue de civilisation.

c) le *bulgare* [2] employé par environ 3.200.000 individus dans le royaume de Bulgarie; il n'a pris qu'au XIXe s. la valeur d'une langue nationale de civilisation.

En Macédoine, la limite entre les parlers de type serbe et de type bulgare ne peut être fixée avec précision. Les parlers macédoniens sont d'ailleurs très évolués.

B. Le *slave occidental* comprend surtout le tchéco-slovaque et le polonais.

Sous le nom de *tchéco-slovaque* on embrasse à la fois les parlers tchèques, employés par 7 à 8 millions d'individus dans l'ancien royaume de Bohême et en Moravie, et les parlers slovaques employés en Slovaquie par environ 2 millions 1/2 d'individus. Comme langue littéraire, le *tchèque* a une importance et une antiquité respectables; les premiers textes littéraires commencent au XIIIe s. et se continuent jusqu'à nos jours. Il y a eu au XIXe s. une renaissance, qui a déterminé la fixation de la langue littéraire liée d'ailleurs à un épanouissement du sentiment national. Il y avait avant la guerre des colonies tchèques abondantes en Autriche (notamment à Vienne) et dans l'Amérique du Nord.

Le *polonais* [3], qui est après le russe la plus importante des langues slaves, est parlé en Europe par 20 à 22 millions d'individus, dont 12 millions dans l'ancienne Pologne russe, 4 millions 1/2 en Galicie, 1.200.000 en Posnanie, 1.300.000 en Silésie,

1. Rešetar, *Die serbo-kroatischen Kolonien Süditaliens*, *ibid.*, vol. IX. Wien, 1911.

2. L. Miletič, *Das Ostbulgarische*, *ibid.*, Wien, 1903.

3. C. Nitsch, *Djalekty języka polskiego*, dans la *Encyklopedja Polska* (en polonais), vol. III, Cracovie, 1915. Le vol. II contient une histoire de la langue.

900.000 dans les deux provinces de Prusse [1], etc. Il faut joindre à ces chiffres 3 millions de Polonais dans l'Amérique du Nord (dont 315.000 à Chicago seulement) et environ 100.000 dans l'Amérique du Sud, notamment au Brésil (état de Parana). On a des textes polonais surtout à partir du XIVe siècle. Il s'est développé dans les derniers siècles une littérature polonaise riche et originale. Les parlers polonais se ramènent aujourd'hui à quatre groupes principaux : le mazovien, le posnanien, le cracovien et le ruthénien.

Le slave occidental s'étendait jadis bien loin vers l'Ouest, jusqu'au delà de l'Elbe; mais il a été peu à peu refoulé et absorbé par l'allemand [2]. On donne parfois le nom de lékhites à l'ensemble des dialectes slaves (y compris le polonais) parlés sur ce domaine. Les dialectes lékhites de l'Allemagne actuelle ne sont guère parlés par plus de 90.000 individus. Le principal est le *sorabe* ou *wende* de Lusace, parlé sur le cours supérieur de la Sprée, dans la région comprise à peu près entre Bautzen et Cottbus [3]. Le *polabe*, très proche parent du wende, qui se parlait sur le cours moyen de l'Elbe dans la région dont Lüchow est le centre, et dont on possède des vocabulaires et de petits textes, est mort au XVIIIe siècle. Plus au Nord, en Poméranie, le *slovince* achève de mourir. Enfin, il faut citer le *kachoub*, encore parlé sur la côte à l'Ouest de Dantzig, et qui est étroitement apparenté au polonais, au domaine duquel il confine.

C. Le *slave oriental* comprend trois dialectes principaux dont la séparation n'est guère antérieure au XIIe siècle :

a) le *grand-russe* [4], dont le représentant le plus important,

1. Paul Stade, *Das Deutschtum gegenüber den Polen in Ost- und Westpreussen*. Berlin, 1908.

2. Behaghel, *Geschichte der deutschen Sprache*, 3e éd. (1911), p. 19 et ss.

3. R. Andree, *Das Sprachgebiet der Lausitzer Wenden vom XVI. Jahr. bis zur Gegenwart*, Prag, 1873.

4. D. Aïtov, *Peuples et langues de la Russie*, *Annales de géographie*, vol. XV (1909), p. 92 ; E. Budde, *Esquisse d'une histoire du russe littéraire contemporain*, XVIIe-XIXe siècles (en russe), dans l'*Enciklopedija slavjanskoj filologij*, t. XII (1908) ; P. Boyer et N. Spéransky, *Manuel pour l'étude de la langue russe*, Paris, 1905.

le dialecte moscovite, a servi de base à la langue russe commune, littéraire et officielle, telle qu'elle s'est fixée au XVIIIe siècle. Cette langue a subi fortement l'influence savante du vieux-slave ecclésiastique. Le grand-russe est une langue de conquérants, qui par suite de la colonisation s'est étendue et imposée à des populations non russes d'origine (surtout finno-ougriennes et tatares). D'après des évaluations qui remontent à l'année 1897, le grand-russe n'était la langue maternelle que de 43 °/o des habitants de l'empire des tsars, soit d'environ 70 millions d'individus; mais l'administration, soutenue par l'école, en a étendu l'usage d'un bout à l'autre de la Russie et en Asie jusqu'aux extrémités de la Sibérie. Il doit y avoir au monde plus de 100 millions d'individus qui parlent russe.

b) le *blanc-russe* (biélorusse) est parlé aujourd'hui dans les gouvernements de Mohilev, Smolensk, Vitebsk et Minsk, c'est-à-dire dans la région qui formait au XIe s. de notre ère ce qu'on appelait la Russie-blanche, qui fut ensuite rattachée au grand-duché de Lituanie et qui passa sous la domination russe en 1772 lors du premier partage de la Pologne. Il s'étend même à une bonne partie des anciens gouvernements de Grodno et de Vilna. En tout, près de 6 millions d'individus parlent le blanc-russe. Toutefois il faut se garder de croire que ce nom désigne une langue commune ou littéraire; le blanc-russe n'est qu'un ensemble de parlers locaux.

c) le *petit-russe* ou *ruthène* [1] s'étend sur un domaine considérable, où il comporte une douzaine de variétés dialectales. Le centre du domaine ruthène est en Russie dans les gouvernements de Poltava, de Kiev, de Kharkov et de Tchernigov, dont l'ensemble constitue ce qu'on appelle l'Ukraine (« Marche »). Mais il s'étend en outre vers le Sud sur les gouvernements de Kherson et d'Ekatérinoslav, où la limite avec le grand-russe est malaisée à fixer. Vers l'Ouest, il embrasse la Podolie, la Volhynie et la Podlachie,

1. A. Meillet, *Le petit-russe et le grand-russe*, dans le *Monde Slave*, 1917 ; Smal-Stocky et Gartner, *Grammatik der ruthenischen (ukrainischen) Sprache*, Wien, 1913.

où d'ailleurs une minorité parle polonais; enfin sur l'ancien territoire austro-hongrois, le ruthène est parlé en Bukovine, en Galicie Orientale et au Nord de la Hongrie, où il confine au polonais, au hongrois et au slovaque [1]. Il y a en tout plus de 30 millions de Ruthènes, dont 4 millions dans l'ancienne Autriche-Hongrie, sans parler de colonies importantes en Amérique (États-Unis et Canada).

J. VENDRYES.

1. O. Broch, *Weitere Studien von der slowakisch-kleinrussischen Sprachgrenze im östlichen Ungarn*, Christiania, 1899.

LANGUES CHAMITO-SÉMITIQUES

Généralités.

Situation ancienne et moderne. — Le chamito-sémitique comprend le sémitique, l'égyptien, le libyco-berbère et le couchitique ; il couvre un vaste domaine continu et peu découpé de contours, dont les limites paraissent avoir peu varié depuis les débuts de l'époque historique : il s'étend d'une part sur l'Arabie et sur les pays qui l'avoisinent au Nord, d'autre part sur la plus grande partie de l'Afrique du Nord, dans toute sa largeur.

La densité de ce domaine linguistique est faible ; ses quelque 20 millions de kilomètres carrés sont coupés de vastes déserts ; aussi le nombre des gens qui parlent sémitique ou chamitique n'excède-t-il pas 50 millions environ.

Les races de ces gens semblent constituer, comme leurs langues, un ensemble relativement cohérent ; ce sont des blancs, légèrement mélangés de nègres au Sud-Est. Leurs civilisations, inégalement développées, suivant les régions et les époques, donnent aussi l'impression d'une certaine unité d'ensemble.

La planche 2 A montre la distribution approximative, en partie hypothétique, du chamito-sémitique vers le 5e s. av. J.-C. : les quatre grands groupes ont leurs domaines séparés et juxtaposés ; le sémitique est subdivisé en multiples langues de civilisation, dont le phénicien seul est à expansion lointaine.

La planche 2 B montre la situation moderne : le sémitique a envahi le chamitique.

L'arabe, la dernière venue parmi les langues littéraires sémitiques, a véhiculé sur un immense espace la civilisation musulmane (chez environ 250 millions d'hommes). Comme langue

religieuse, il s'étend en Asie en dehors des limites de cette carte; le vocabulaire arabe a pénétré en masse dans le lexique de langues importantes.

Les autres langues sémitiques ont péri en majorité, l'arabe occupant maintenant leurs anciens domaines : il reste de petites régions araméennes et sudarabiques, surtout en bordure du territoire de l'arabe. L'hébreu n'a plus de territoire propre, mais il est resté langue religieuse et savante chez les juifs dispersés par le monde. Seul le groupe des langues éthiopiennes, colonie sémitique pleine de vitalité, continue à s'étendre en Afrique orientale.

Les langues chamitiques ont été en grande partie recouvertes par le sémitique. L'égyptien ne survit que dans l'usage liturgique du copte qui n'est plus parlé ni écrit. Les dialectes berbères coupés en îles et îlots, reculant en beaucoup de points encore de nos jours devant l'arabe, très rarement écrits, semblent avoir perdu toute chance de former une langue de civilisation. Le couchitique a été presque éliminé du haut plateau éthiopien par le sémitique ; mais il reste vivant sur ses confins ; notamment au Sud, les langues galla et somali, quoique non écrites, ont un rôle appréciable.

Au XIX[e] siècle, des langues européennes ont pris pied dans la région méditerranéenne de l'Afrique, avec des apports de colons.

Divisions ; relations avec d'autres groupes. — La grammaire comparée du chamito-sémitique n'est pas faite : certaines concordances morphologiques claires et la grande ressemblance des systèmes phonétiques permettent d'affirmer la parenté entre eux des quatre groupes considérés, malgré les grandes différences qui les séparent. Une étude approfondie du vocabulaire permettrait seule d'établir des correspondances phonétiques précises et de juger correctement certains éléments morphologiques ; elle reste presque toute à faire.

L'étude intérieure des groupes est malheureusement elle-même trop peu avancée. Seules les langues sémitiques ont été bien examinées : grandes langues de civilisation, attestées sur une longue période par des textes, presque toutes s'imposaient à l'étude ;

l'évidence de leur parenté incitait aux rapprochements. Toutefois la grammaire comparée du sémitique est loin d'avoir atteint encore le degré de perfection de la grammaire comparée des langues indo-européennes. La grammaire de l'égyptien ancien est encore assez mal connue ; embarrassés par une écriture compliquée, insuffisamment guidés par des comparaisons avec d'autres langues, les savants ne s'y sentent pas encore toujours sur un terrain sûr. Pour le libyco-berbère, si les dialectes modernes sont d'année en année mieux connus, la langue ancienne, très mal attestée, échappe encore à l'étude. Les dialectes couchitiques, nombreux, sont mal explorés, et tous à l'époque moderne. Aussi le présent chapitre ne peut-il être qu'une mise en place provisoire où plus d'une fois l'énumération doit remplacer le classement.

En particulier, la division en quatre groupes séparés paraît seule prudente dans l'état présent des recherches. Le rapprochement étroit que certains ont voulu établir entre l'égyptien et le sémitique n'est aucunement prouvé. Il n'y a pas lieu non plus de croire à la parenté spéciale entre l'égyptien, le libyco-berbère et le couchitique que suppose leur réunion habituelle sous le nom de chamitique ; il ne sera donc pas question ici d'un groupe chamitique.

Le terme de sémitique a été adopté à la fin du XVIII[e] siècle par les savants européens parce que les peuples parlant les langues sémitiques sont en majorité compris parmi la postérité de Sem (en hébreu *šem*) dans le chapitre X de la Genèse. C'est à la même source biblique que, dans la seconde moitié du XIX[e] siècle, on a puisé le terme de chamitique (hamitique, khamitique), d'après hébreu *ḥām*, grec des Septante *kham* « Cham ». Enfin le terme de couchitique (kouchitique, couschite, etc.) a été ensuite fait sur le même modèle d'après le nom de *kūš*, qui désigne dans la Bible celui des fils de Cham dont les descendants semblent situés le plus au Sud ; *kꜥš* dénomme d'autre part en égyptien les pays au Sud de l'Égypte ; le terme de couchitique paraît donc assez bien adapté à la désignation des langues non sémitiques et non soudanaises de la région abyssine ; en conséquence le nom d'éthiopien doit être (en dépit de certains auteurs) réservé aux langues sémitiques

d'Abyssinie, et couchitique ne doit jamais être pris comme synonyme de chamitique. Le nom de protosémitique, au lieu de chamitique, n'a pas fait fortune.

La division en quatre groupes seulement qui est suivie ici exclut l'incorporation au chamito-sémitique, voulue par certains auteurs, du haousa, du peul et du nama (hottentot) qui dans ce livre sont classés ailleurs. De ces trois langues, seul le haousa présente quelques traits isolés d'homonymie avec des éléments grammaticaux chamito-sémitiques.

Il est prématuré de tenter le rapprochement du chamito-sémitique ci-dessus défini avec d'autres familles de langues. Ce n'est pas à dire que les tentatives partielles qui ont été faites jusqu'à présent ne gardent pas pour l'avenir une certaine valeur de travaux d'approche. La plus poussée est celle de M. Möller pour comparer le vocabulaire sémitique au vocabulaire indo-européen. L'idée maîtresse d'un essai dans cette direction est de rechercher une unité ancienne des principales langues parlées par les peuples de race blanche. Au contraire certains savants cherchent à mettre en évidence une unité africaine : Reinisch, surtout, s'est montré attentif aux transitions possibles entre les langues chamitiques et la famille bantou.

Caractéristiques. — Des caractéristiques communes à tous les groupes chamito-sémitiques sont examinées ici brièvement. Les exemples donnés servent en même temps à montrer les principales preuves de la parenté de ces groupes entre eux. Certains des traits communs décrits peuvent n'avoir pas appartenu à l'ancêtre unique chamito-sémitique, mais résulter de développements parallèles : le parallélisme du développement dans les détails est encore un indice de parenté.

La phrase chamito-sémitique est composée de mots nettement séparés, en général pourvus d'un accent distinct, dont les principaux sont des verbes ou des noms. Chaque mot reçoit toutes les caractéristiques nécessaires soit pour indiquer les modifications secondaires de la notion principale qu'il exprime, soit pour marquer son rôle dans la phrase ; il est indépendant de l'aspect phonétique de ses voisins.

Le court exemple suivant est pris à l'arabe classique, la langue chamito-sémitique la plus abondante en signes de flexion.

(Coran XII, 46) *yūsufu 'ayyuhā (l) ssiddīqu, 'aftinā fī sabʿi baqarātin simānin ya'kuluhunna sabʿun ʿidžāfun.*

yūsufu = Joseph (*-u* = marque du cas sujet ou vocatif)

'ayyuhā = ô (adresse solennelle)

-(l)s-siddīqu = le (article *(a)l-*, dont *l* est assimilé à la consonne suivante) juste (*-u* comme ci-dessus)

'afti-nā = décide (impératif d'un thème à préfixe *'a-* d'une racine *fty*) -nous (pronom suffixe)

fī = au sujet de (dans)

sabʿi = sept (*-i* est la marque du cas régime de nom ou cas après préposition)

baqarātin = vaches (pluriel de *baqarat-un* ; *-n* est le signe de l'indétermination)

simānin = grasses (pluriel de *samīn*)

ya'kulu-hunna = mange(nt) elles (verbe *'akala* ; *ya-* préfixe de sujet 3ᵉ personne masc. singulier *-hunna* suffixe de régime 3ᵉ pers. plur. ; le verbe n'est pas accordé avec le sujet qui suit et il n'y a pas de relatif exprimé)

sabʿun = sept (cas sujet, indéterminé)

ʿidžāfun = maigres (pluriel de *ʿadžīf* ; *dž* représente une consonne unique)

« Joseph, ô le juste, éclaire-nous sur sept vaches grasses que mangent sept vaches maigres ».

(Voir aussi les phrases citées pp. 96, 105, 130, 135, 142.)

Le centre du mot est une racine. Chaque racine a un certain nombre d'éléments essentiels qui sont généralement des consonnes, mais peuvent être aussi des voyelles longues en alternance avec semi-voyelle (ce qui se rencontre dans une partie du vocabulaire sémitique) ou stables (ce qui est fréquent en couchitique) ; voir en outre les indications de la page 92. Les éléments radicaux sont constants dans les mots formés avec la racine ; ils supportent seuls l'idée exprimée, qu'il s'agisse d'une action ou d'un objet. L'importance particulière de l'armature des racines

est visible, au sens propre, dans les écritures qui ont été inventées pour des langues chamitiques et sémitiques : en principe les consonnes seules y sont notées. Exemple de racine : égyptien et sémitique *mwt*, berbère *mmt*, « idée de mort ».

Le squelette-idée ainsi défini n'est pas le mot, qui seul apparaît en fait dans le langage. Le mot est caractérisé dans un emploi déterminé de verbe, de nom, etc., et ce qui le caractérise le plus souvent, c'est le jeu des voyelles qui s'insèrent entre les consonnes radicales : ainsi l'arabe a *fataḥa* « il a ouvert » *fatḥ(un)* « commencement, conquête », etc. Un pluriel comme le *simān-* (singulier *samīn-*) de la phrase analysée ci-dessus montre comment des modifications vocaliques servent à marquer le rôle des mots. Pour le rôle des affixes, voir pages 87-90.

Les voyelles qui animent les racines consonantiques caractérisent par leurs combinaisons variées tant des formes verbales que diverses catégories de noms : il n'y a pas priorité, en général, du verbe sur le nom ou du nom sur le verbe ; l'un et l'autre peuvent se dériver indépendamment de la plupart des racines ; ainsi une racine telle que l'éthiopien *ngr* est aussi bien un nom radical, comme *nagar* « parole, chose », qu'un verbe *nagara* « il a dit » ; le copte *shaï* est « écrire » et « écriture » ; le berbère du Sous marocain a *mgör* « fauche », *imgör* « faucille ».

Mais ce principe est loin de gouverner seul la constitution du lexique.

Tout d'abord, dans un certain nombre de racines, le nom apparaît comme seul primitif : le verbe, s'il existe, est dérivé ; ainsi arabe *ridẑl(un)* « pied », *radẑul(un)* « homme », hébreu *regel* « pied » montre une racine nominale ; secondairement l'arabe tire du nom du « pied » un verbe dénominatif *radẑdẑala* « il a affermi sur ses pieds ». D'autre part un grand nombre de noms ne se rattachent pas directement à une racine, mais à un thème verbal dont ils ont la vocalisation, les modifications consonantiques internes, les préfixes ; ils sont eux-mêmes dérivés de ce thème au moyen d'une vocalisation spéciale, d'un préfixe (*m-* est le plus fréquent) ou d'un suffixe. Ce ne sont pas seulement des infinitifs ou noms d'action et des participes (pouvant jouer

à l'occasion le rôle d'adjectif ou de substantif), mais encore des noms de lieu et d'instrument. Des exemples seront donnés ci-dessous avec les thèmes verbaux.

Un verbe chamito-sémitique est en effet l'ensemble d'une série de thèmes, eux-mêmes quelquefois pourvus de différentes vocalisations ; les aspects différents de la racine expriment des modes différents de l'idée. Le verbe ainsi équipé joue un grand rôle dans les langues chamito-sémitiques, malgré l'emploi répandu de la phrase nominale pour exprimer les situations et les qualités.

Des nuances vocaliques qui permettent de distinguer au thème le plus simple un verbe actif d'un verbe neutre, et à la plupart des thèmes, d'exprimer l'opposition de l'actif et du passif, ne sont attestées que dans certaines langues sémitiques : ainsi arabe *qatala* « il a tué », *qutila* « il a été tué » ; *labisa* « se revêtir », mais *labasa* « couvrir ».

Plus répandus sont les procédés de renforcement interne de la racine pour l'expression de divers renforcements de l'idée. On peut noter comme principaux types : la gémination de consonne médiane (ainsi arabe *qattala* « tuer net », *qattāl* « grand tueur, tueur de métier ») ; la répétition de la première partie du mot (ainsi somali *dabdabar* « enchaîner en rang » de *dabar* « enchaîner ») ; l'insertion d'un *ā* d'après la première radicale (ainsi bedja *dālib* « commercer », en face de *delib* « conclure un achat ou une vente » ; arabe, avec nuance conative, *qātala* « combattre, chercher à tuer », *muqātil(un)* « qui combat »).

Mais le principal procédé de formation de thèmes secondaires est l'application de préfixes aux thèmes simples et renforcés ; ces préfixes peuvent être eux-mêmes complexes. Les mieux attestés dans les différents groupes de langues sont le *t* réfléchi, qui exprime souvent le passif (il est quelquefois infixé) et le *s* ou *š* causatif :

arabe *tafarraqa* et *(i)ftaraqa* « il a été séparé » (racine *frq* « fendre ») ; de *tafarraqa*, les noms d'action *tafarruq(un)*, *mutafarraq(un)* ; hébreu *hitpārǒqū(w)* « ils ont été arrachés » (même racine, avec *p* correspondant à *f*) ; berbère marocain (*Ait Sögruššön*) *tuăšör* « sois volé » (*ăšör* « vole ») ; bedja *tōbās* « être enterré » (de *bis* « enterrer ») ;

amharique *asărrama* « il a fait sarcler, il a fait corriger » (du simple *arrama*), infinitif *măsărram* « action de faire sarcler », nom d'instrument *măsărrămyă* « ce qui sert à faire sarcler, ce qui concerne l'acte de faire corriger »; arabe (*i*)*staḫbara* « il s'est informé », d'une racine *ḫbr* « apprendre, être informé » ; accadien *ušakšad* « il fait conquérir », de *kšd* « conquérir » ; égyptien *śnfr* « rendre beau » de *nfr* « être beau » ; berbère (kabyle) *sökšöm* « fais entrer », en face de *ökšöm* « entre » ; afar *sgafa* « faire tuer », de *gafa* « tuer ».

Si des procédés analogues peuvent produire, au départ d'une racine, soit des thèmes verbaux, soit des thèmes nominaux, la distinction tranchée du verbe et du nom apparaît dans la flexion.

Le verbe, presque toujours, est nettement distingué par des désinences personnelles. Les désinences expriment à elles seules le sujet de première ou deuxième personne (arabe *'a-ktubu* « j'écris »); à la troisième personne elles peuvent aussi désigner à elles seules un sujet connu par ailleurs (*ya-ktubu* « il écrit »); elles existent, inséparables du verbe, même si le sujet est exprimé par un nom ou un pronom (*ya-ktubu rradžulu* « (il) écrit l'homme ». Les désinences personnelles comportent des distinctions de genre (masculin-féminin, aux 2ᵉ et 3ᵉ personnes seulement) et de nombre (singulier, duel, pluriel).

Les seules caractéristiques qui paraissent conservées de l'unité chamito-sémitique sont les caractéristiques préfixées (qui sont complétées au moyen de suffixes exprimant certaines des distinctions de genre et de nombre). En sémitique occidental elles ne servent que pour la forme de l'action inaccomplie (imparfait), par opposition à la forme de l'action accomplie (parfait) qui est pourvue de suffixes (*katabta* « tu as écrit »). Mais en sémitique oriental, en berbère et en couchitique (l'égyptien est à part) il n'y a que des formes à préfixes : l'opposition des deux états de l'action ou bien se fait au moyen de différences vocaliques, ou bien fait défaut. Il semble probable que la distinction des deux aspects est secondaire, et que le chamito-sémitique n'avait à l'indicatif qu'une forme conjuguée, exprimant l'action sans en

distinguer les états. Le temps à trois divisions — passé, présent, futur — n'a pas d'expression morphologique.

Les exemples suivants montrent la caractéristique préfixée *t-* dans son double emploi de 2[e] personne (ici au masculin singulier) et de 3[e] personne féminin singulier : arabe *taqtulu* « tu tue(ra)s, elle tue(ra) » ; accadien *takašad* « tu conquiers (conquerras), elle conquiert (conquerra) », *takšud* « tu as, elle a conquis » ; berbère marocain (*Ait Sögruššön*) *þgörsöt* « tu égorges, tu as égorgé » *þgörs* « elle égorge, elle a égorgé » ; afar *taggifä* « tu tue(ra)s, elle tue(ra) », *tiggifä* « tu as, elle a tué ».

La flexion du nom comporte des indications de genre, de nombre, de cas qui ne sont complètes que sur certains points du sémitique (arabe, ancien accadien). Rien ne peut en être attribué avec une suffisante vraisemblance au chamito-sémitique commun, si ce n'est le *t* qui est très souvent la marque du féminin, et l'état construit.

Le *t* féminin est visible dans les exemples suivants. Il ne sert pas seulement au rôle propre de féminin, c'est-à-dire à la distinction des sexes (surtout dans les adjectifs), il exprime aussi dans nombre de cas la restriction (diminutif) et l'abstraction. Exemples : guèze *bö'sīt* « femme » (de *bö'si* « homme), *söqlat* « crucifixion », de racine *sql* « pendre » ; arabe moderne tunisien (à Djendouba) *ḥagra* (le *t* non apparent dans cette forme se révèlerait, à la place de *a*, si on mettait le mot à l'état construit) « marque de dédain » en face de *ḥǫgǫr* « dédain » ; vieil égyptien *s*ᶜ*t* « fille » (*s*ᶜ « fils ») ; berbère marocain (Beni Mgild) *timiššūt* « chatte » (*imiššu* « chat »).

L'état construit est le terme qui désigne au sens étroit la forme spéciale que prend en sémitique un nom immédiatement suivi de son complément ; on peut en étendre l'emploi à l'ensemble des phénomènes qui caractérisent le rapport d'un nom avec un complément qui s'y annexe sans préposition interposée. Le fait essentiel est celui-ci : les deux mots rapprochés, régent et régi, forment un seul groupe accentué dont les termes composants n'ont qu'une autonomie réduite ; les groupes ainsi formés sont des composés occasionnels. Les effets de cette sorte de composition sur les deux termes sont variables : les deux termes peuvent

n'être pas modifiés, en dehors du changement d'accentuation ; mais souvent ils le sont : le premier terme (régent) peut avoir une finale autre qu'à l'état isolé, il peut aussi changer son aspect vocalique interne ; le second terme (régi) peut être changé dans son initiale :

hébreu *su(w)se(y) parʿōh* « les chevaux (à l'état indépendant *su(w)si(y)m*) du Pharaon » ; *dǝbar 'ĕlōhi(y)m* « la parole (état indépendant *dåbår*) de Dieu » ; arabe moderne libanais *baqrt el-föllāḥ* « la vache (état indépendant *baqra*) du paysan » ; guèze *qāla 'ǝgzi'abǝḥēr* « la voix (état indépendant *qāl*) de Dieu », *gětā mǝdr* « le seigneur (état indépendant : même forme) de la terre » ; copte *gǝb ẓoyit* « feuille (état indépendant *goobe*) d'olivier » ; berbère (kabyle) *aman ugǝlmim* « l'eau de l'étang (état indépendant : *agǝlmim*) » ; agaw (kemant) *wāẓin lānqā* « flamme (2) de feu (1) » ; on remarquera l'ordre des mots propre au couchitique (les deux termes sont pareils à l'état indépendant).

Une autre annexion de compléments, qui n'est pas sans rapport avec l'état construit, et n'est pas moins caractéristique du chamito-sémitique, est ce qu'on pourrait appeler la flexion pronominale des verbes et des noms. Un complément pronominal, qui désigne l'objet avec le verbe et le possesseur avec le nom, s'adjoint au terme complété comme une véritable désinence variable en personne, en genre (aux 2ᵉ et 3ᵉ personnes) et en nombre. Ces désinences s'emploient aussi en suffixes avec les prépositions ; elles sont un peu différentes avec le verbe d'une part, le nom et la préposition d'autre part ; ces éléments subissent par l'adjonction des suffixes des modifications analogues à celles de l'état construit. Les pronoms personnels indépendants ont en partie les mêmes formes que les pronoms suffixes, mais allongées ou adjointes à d'autres éléments pronominaux :

hébreu *dǝbārǝ-xā* « ta parole » ; guèze *qālǝ-ka* « ta voix » ; égyptien *pr-k* « ta maison » ; berbère (kabyle) *aḫḫam-ik* « ta maison » ; saho-assaorta (le pronom est préfixé, non suffixé) *kubaglā* « ta mule ».

L'aperçu rapide sur le chamito-sémitique serait incomplet s'il

n'était pas tenu compte du système d'articulation. L'originalité de ce système n'est pas seulement la variété et la solidité des consonnes, mais aussi leur qualité ; le fait le plus notable est l'abondance des consonnes glottales et laryngales. Il est frappant que le bruit de séparation brusque des cordes vocales (occlusive glottale, coup de glotte, attaque vocalique brusque) soit compté comme une consonne (notée ʾ). Il s'y ajoute une spirante glottale qui est le bruit de la voix chuchotée (*ḥ*) ; une spirante laryngale obtenue par une constriction du larynx (c'est la voix pressée qu'on fait émettre aux patients pour montrer leur gorge au médecin), notée ʿ ; un *h* nettement articulé ; une ou deux spirantes arrière-vélaires (*ḫ*, *ġ*). De plus les consonnes dites emphatiques (*q*, *ṭ*, *ṣ*, etc.) comportent, plus encore qu'une tension générale des organes d'émission, une contraction de la région laryngale. Les consonnes de l'avant de la bouche (surtout les dentales, et les prépalatales sur certains domaines) sont aussi abondamment représentées. Mais les labiales sont en petit nombre.

Sémitique.

Division ; caractéristique. — La division généralement admise des langues sémitiques connues est :

Sémitique oriental : accadien.
Sémitique occidental : *a*) septentrional : cananéen, araméen.
b) méridional : arabe, sudarabique, éthiopien.

Les langues sémitiques ont une caractéristique commune, la trilitéralité ; elle consiste en ceci que presque toutes les racines sont composées de trois consonnes ; les deux dernières de ces consonnes peuvent être identiques ; l'une des consonnes peut être une semi-voyelle qui, dans certaines formes, apparaît comme une voyelle longue. Ainsi racines *qtl* « tuer », *tmm* « être com-

plet », *qwm* - *qūm* « se tenir debout » (en arabe *ya-qūm-u* « il se tient debout », *qawm(un)* « action de se tenir à un endroit »). Il est assez probable que la trilitéralité générale ne date pas de la période chamito-sémitique. En effet, en sémitique même, certaines racines parentes par le sens ont en commun deux consonnes seulement. D'autre part les groupes apparentés, ou bien ont en abondance (égyptien et surtout berbère) des racines à deux consonnes, ou bien ont normalement (couchitique) un type à deux consonnes encadrant une voyelle qui n'alterne pas avec semi-voyelle (bilin : *bän* « partager »). Le doute sur l'origine fait qu'on peut considérer la trilitéralité généralisée comme d'autant plus caractéristique du type sémitique. Elle s'atténue dans des langues où ce type est altéré (ainsi en amharique); elle persiste comme tendance agissante dans une langue conservatrice comme l'arabe (où par exemple un bilitère ancien *yad(un)* « main » a passé à *yidd* en maghribin moderne).

Sémitique oriental.

Le sémitique oriental est défini morphologiquement par sa conjugaison qui comprend deux temps à préfixes (accompli : *ikšudū* « ils ont conquis », inaccompli : *ikašadū* « ils conquièrent ou conquerront ») et un temps à suffixes exprimant la durée, dit permansif (*kašdū* « ils sont (étaient, seront) en train de conquérir »). Le consonantisme est altéré : la plupart des laryngales manquent.

On ne connaît jusqu'à présent qu'une langue sémitique orientale : l'accadien. Mais des découvertes ultérieures en exhumeront peut-être d'autres (sur le classement douteux de l'amorite, voir p. 98 et à l'index).

Histoire de l'accadien. — L'accadien est à l'origine la langue de Sémites qui ont envahi un pays de langue et de civilisation sumériennes. Le sumérien a subsisté sous la domination accadienne, surtout comme langue religieuse, littéraire et officielle : mais il a dû céder le rôle principal à l'accadien qui a eu une longue fortune comme langue de civilisation de puissants empires.

Le territoire central de ces empires couvrait environ 250.000 kilomètres carrés ; mais à maintes périodes ils ont plus que doublé l'étendue de leur domination. Les parties fertiles en devaient être assez peuplées.

Quant à la langue, si les monuments la montrent un peu différente suivant les périodes et les lieux, elle est pourtant foncièrement une.

La chronologie de l'histoire suméro-accadienne n'est pas encore établie sans conteste, surtout pour la période la plus ancienne que la chronologie « longue » fait débuter entre 3800 et 3700 et que la chronologie « courte » date au plus de 3000 av. J.-C. Pour l'époque encore antérieure, des listes royales mal situées dans le temps donnent des indications qui sont sans doute à répartir sur plusieurs siècles.

La plus ancienne période historique, qui se prolonge jusqu'après 2000, est celle où prédominent des États situés dans la Mésopotamie entre le Tigre moyen et l'Euphrate moyen et dans la région voisine jusqu'au Golfe Persique. Ils ont eu leur centre en Chaldée d'abord au Nord, dans le pays d'Akkad (désignation indigène) ; puis, au Sud, dans l'ancien pays de Sumer (*šumeri* en désignation indigène) ; enfin de nouveau au Nord, avec Babylone comme capitale. — La littérature de cette période a été conservée jusqu'à nous dans les sables du pays : inscriptions monumentales sur les statues, livres et lettres écrits sur des briques ont livré aux déchiffreurs l'« ancien accadien » ou « (ancien) babylonien » (qui est caractérisé linguistiquement par la déclinaison nominale à trois cas avec *-m* final).

Dans le second millénaire, la Babylonie cesse d'exercer l'hégémonie. Souvent dévastée, gouvernée pendant de longues périodes par des étrangers, elle finit par être soumise à ses voisins du Nord.

Un État sémitique a existé dans la région montagneuse de la vallée du Tigre depuis 2400 environ (documents du vieil assyrien). Au cours du second millénaire, des rois qui avaient leur capitale à Assur (*Aššur*), puis à Ninive, se révèlent grands conquérants et disputent l'hégémonie de l'Asie antérieure à de

puissants voisins ; leur empire, après avoir hérité de la puissance babylonienne en décadence, a sombré définitivement en 606 av. J.-C. On a retrouvé sur le domaine septentrional de nombreux documents dont on désigne la langue par le nom d'assyrien.

En dernier lieu, un nouvel empire méridional a fleuri en Babylonie de 626 à 539 (date de la prise de Babylone par les Perses de Cyrus). La langue de cet empire est le « néo-babylonien ».

Aux périodes de plus grande extension, l'accadien avait été employé au loin en dehors même des limites des empires babylonien et assyrien. Ainsi on a retrouvé à Tell el-Amarna en Égypte une correspondance en accadien échangée vers 1400 av. J.-C. entre des souverains égyptiens et des princes orientaux, notamment des cananéens. L'Asie Mineure livre aussi aux fouilleurs des documents accadiens, comme les tablettes qui sont dites « cappadociennes » du lieu de leur découverte.

Après la chute de Babylone, la langue accadienne n'a pas persisté beaucoup dans l'usage parlé ; depuis le IVe s. av. J.-C. environ elle n'a plus été qu'une langue savante et religieuse, qui a dû être utilisée jusqu'aux environs de l'ère chrétienne. C'est l'araméen qui s'est substitué à l'accadien comme langue parlée et comme langue diplomatique.

Le terme « accadien » employé ici, déjà proposé il y a longtemps par Oppert, paraît devoir se stabiliser comme désignation des sémites et des dialectes sémitiques du domaine ci-dessus défini ; il remplace avantageusement assyrien et babylonien, désormais réservés à des désignations partielles, et le terme compliqué d'assyro-babylonien. Malheureusement, quand on a commencé à étudier le sumérien on l'a d'abord quelquefois nommé accadien. D'autre part le terme « assyriologie » pour l'étude de tout ce domaine est si bien implanté qu'il ne paraît pas devoir disparaître. Les assyriologues n'emploient jamais le mot « chaldéen » pour désigner un dialecte accadien ; mais ce nom est appliqué aux faits de civilisation et d'histoire accadienne méridionale (art chaldéen, empire néo-babylonien ou chaldéen).

Écriture cunéiforme. — L'accadien a été écrit dans le système « cunéiforme » dont usaient les Sumériens et qui a été appliqué à des langues variées de l'Asie occidentale. C'était à l'origine une notation par idéogrammes, dessins représentant des objets. Les idéogrammes linéaires schématiques sumériens, écrits en cursive sur l'argile molle, sont devenus des combinaisons de « clous » et de « coins », d'où le nom descriptif de « cunéiformes ». Les caractères cunéiformes ont été peu à peu employés par les Sumériens comme signes syllabiques en même temps que comme idéogrammes (ainsi l'étoile qui se lisait *ana* « ciel » ou *dingir* « dieu » était aussi la syllabe *an*). Les Accadiens ont conservé les valeurs des idéogrammes, mais les lisaient en sémitique, et ont d'autre part maintenu des valeurs syllabiques sumériennes ; l'étoile est lue *šamû* « ciel », *ilu* « dieu », et *an*. De plus les mots sémitiques à leur tour peuvent fournir des valeurs syllabiques : ainsi un signe représentant une main (qui avait en sumérien la valeur *šu*), est lu *qatu* « main » ou, comme syllabe, *qat*.

Les syllabes notées sont de toutes formes, ainsi *a*, *na*, *ab*, *nab* ; une syllabe à trois éléments comme *nab* peut être à l'occasion écrite *na-ab*, ou *na-a-ab* (pour noter expressément un *a* long).

Si donc l'écriture cunéiforme a l'avantage de noter les voyelles et pas seulement le squelette consonantique des mots, elle a l'inconvénient d'une très grande complication. Pour lire un texte accadien il faut connaître environ 300 signes dont la plupart ont plusieurs valeurs ; en outre les formes varient avec les périodes. Aussi après environ trois quarts de siècle de travail depuis le début du déchiffrement de l'accadien, la lecture des nouveaux textes mis au jour offre-t-elle encore quelquefois des incertitudes.

Les caractères cunéiformes sont séparés ; ils se lisent en général horizontalement de gauche à droite (anciennement, de haut en bas). L'exemple suivant montre le mélange de syllabes détachées, d'idéogrammes (représentés par des mots écrits sans divisions) et d'idéogrammes « indicatifs », mis ici entre parenthèses, qui n'étaient sans doute pas prononcés (termes géographiques, noms des dieux, des rois, etc.).

Annales de Tukulti-Ninip II, lignes 49-50 :

a-na eli	(*nār*)	*Diglat*	*aq-ṭi-rib-ma*	*maš-ka-na-a-te*
vers	[le] fleuve	Tigre	je m'approchai-et	[les] villages
ša (*māt*)	*U-tu-'a-a-te*	*al*	*qab-ra-ni-šu-nu*	*ša*
du pays	Utuate	ville	[de] leurs tombes	qui
šit-ku-nu	*eli*	(*nār*) *Diglat*	*ak-ta-šad.*	
est située	sur	[le] fleuve Tigre	je conquis.	

Sémitique occidental.

Les langues sémitiques occidentales opposent dans le verbe un parfait pourvu de suffixes à l'imparfait pourvu de préfixes (et suffixes), et ignorent la forme durative (permansif).

Imparfait : hébreu *yiqṭölū*, araméen *yiqṭölūn* (et *nöqṭölūn*), arabe *yaqtulū*(*na*), guèze *yöqattölū* « ils tue(ro)nt ».

Parfait : hébreu *qāṭölū*, araméen *qöṭal*(*ū*), arabe *qatalū*, guèze *qatalū* « ils ont tué ».

(Pour l'hébreu voir en outre ci-dessous p. 105.)

Alphabet sémitique. — C'est sur le domaine occidental que s'est développée l'écriture sémitique. En chamito-sémitique cette écriture a été utilisée en outre par le libyco-berbère. D'autre part elle a couvert l'Europe et une partie de l'Asie ; l'expansion européenne la véhicule de nos jours dans le monde entier sous la forme de l'alphabet latin.

La question de son origine n'est pas encore bien résolue. L'absence de documents de cette écriture dans les correspondances de Tell el Amarna (vers 1400 av. J.-C.) semble indiquer qu'elle n'était pas encore répandue à cette époque. En fait, jusqu'à la fin du XIX[e] s. on n'en connaissait aucun monument plus vieux que les abords de 1000 av. J.-C. ; on avait cherché en vain à la rattacher à des systèmes plus anciennement attestés tels que les hiéroglyphes égyptiens et les écritures égéennes. Dans les premières années du XX[e] s. on a trouvé dans la presqu'île du Sinaï, sur le site appelé Serabit el-Khadim, à côté d'inscriptions hiéro-

glyphiques anciennes, onze courtes inscriptions non hiéroglyphiques. Ces inscriptions contiennent en tout environ 150 caractères qui se laissent classer en un nombre de signes aussi réduit que celui des lettres sémitiques. Certains de ces signes ressemblent à des hiéroglyphes, d'autres à des signes phéniciens ; la direction serait de haut en bas et de gauche à droite (aussi de droite à gauche), d'après le premier déchiffrement, qui livre au moins un mot sémitique. La date de ces monuments serait, de manière large, comprise entre 2000 et 1500 av. J.-C. Si on considère les présomptions fournies par ce « paléosinaïtique » sémitique, on est amené à penser que les inventeurs de l'écriture sémitique avaient une connaissance au moins extérieure du système égyptien.

Même si les inventeurs de l'écriture sémitique occidentale ont en partie imité l'écriture égyptienne, leur originalité reste grande. En effet les Egyptiens n'ont jamais su isoler de la masse de leurs autres signes les signes phonétiques dont ils se servaient. D'autre part les systèmes syllabiques connus ailleurs dans le monde antique (par exemple la partie phonétique du système accadien) ont toujours des consonnes accompagnées d'une voyelle déterminée. Au contraire, ce que le système sémitique montre aux yeux, ce sont essentiellement les consonnes, qu'on peut lire seules ou avec une voyelle quelconque.

L'analyse phonétique sémitique a permis de réduire à un très petit nombre le total des signes ; mais, encore incomplète, elle a l'inconvénient de ne pas spécifier les voyelles de chaque syllabe.

Ce manque de signes distincts pour les voyelles a relativement peu d'inconvénient pour certaines langues sémitiques où le solide squelette consonantique suffit souvent à suggérer non seulement la racine portant la signification, mais le thème verbal ou nominal et même les désinences : un lecteur lisant sa langue maternelle n'a que relativement peu de peine à suppléer les voyelles caractéristiques de chaque forme, constantes pour une formation donnée.

Pourtant les Sémites ont été gênés par l'absence de voyelles distinctes dans l'écriture, surtout quand ils ont eu à lire des documents d'une langue autre que leur parler maternel, ainsi qu'il

arrive dans le cas de textes liturgiques écrits soit dans une langue morte, soit dans une forme ancienne de la langue vivante (hébreu, arabe coranique, etc.). La difficulté a été résolue quelquefois partiellement par l'emploi de certaines consonnes laryngales affaiblies dans la prononciation et des semi-voyelles *w* et *y* pour noter au moins les voyelles longues. Sur certains points il a été fait usage de signes complémentaires sur ou sous les consonnes pour noter même les voyelles brèves. L'éthiopien a constitué un syllabaire en accrochant aux consonnes des marques pour signifier les voyelles.

A cause de l'absence de notation ancienne des voyelles, on est mal renseigné sur bien des langues qui ne sont attestées qu'en squelette consonantique (comme le phénicien) ou dont le vocalisme a été noté après qu'elles avaient disparu de l'usage parlé (comme l'hébreu). Des détails complémentaires seront donnés à propos de chaque langue.

La direction ordinaire de l'alphabet sémitique occidental est la ligne horizontale de droite à gauche ; dans la suite, seules les déviations de cette direction seront indiquées.

SÉMITIQUE (OCCIDENTAL) SEPTENTRIONAL

Il comprend principalement le cananéen et l'araméen, proches entre eux. Peut-être faudra-t-il y ajouter le « (paléo)sinaïtique » (voir p. 97). Peut-être aussi y joindra-t-on l' « amorite » (amoraïque, amorrhéen). Il s'agit d'un ancien peuple dont l'activité semble s'être étendue de la Syrie à la Mésopotamie vers une période 2800-1500 av. J.-C. (*Amurru* des documents accadiens). On s'attache à restituer quelque chose de sa langue, au moyen de noms propres compris dans les documents en accadien, mais cette recherche rudimentaire ne permet pas de classement solide; il n'est même pas prouvé que l'amorite soit sémitique (voir p. 92.)

Le sémitique septentrional est caractérisé par la réduction du nombre de consonnes (par rapport au sémitique méridional présumé plus conservateur), l'absence de la flexion casuelle et l'usage constant du pluriel à suffixes dans les noms.

Cananéen.

Le terme « cananéen » est tiré du nom hébreu Canaan (*Kōnaᶜan*) qui désignait les régions de Phénicie et Palestine et leurs habitants (soit peuples non sémites anciennement établis, soit envahisseurs sémites installés avant l'arrivée des Hébreux).

Le cananéen, qui comprend comme langues importantes l'hébreu et le phénicien, semble n'avoir jamais couvert qu'une petite région continentale de 80.000 kilomètres carrés environ, à laquelle il faut ajouter les colonies phéniciennes.

Vieux cananéen. — Les plus anciens mots écrits de cananéen qui soient conservés sont des gloses en caractères cunéiformes dans les correspondances de princes cananéens adressées aux rois d'Égypte Aménophis III et Aménophis IV ou Ikhounaton vers 1400 av. J.-C. ; ces lettres font partie des tablettes de *Täll el-ʿamarna* (emplacement de la capitale d'Ikhounaton).

Moabite. — Ce nom désigne la langue d'une seule inscription, due à un roi de Moab (au S.-E. de la mer Morte), *Mešaᶜ* (on écrit généralement Mésa) qui raconte ses démêlés avec le roi d'Israël au milieu du IXᵉ s. av. J.-C. Cette langue est très proche de l'hébreu.

La stèle de Mésa est le plus ancien monument connu de l'écriture cananéenne (ou phénicienne) qui est celle des inscriptions phéniciennes et hébraïques et s'est perpétuée dans celle des manuscrits samaritains. Elle comporte 22 lettres.

Phénicien, punique. — Le phénicien est la langue des nombreuses inscriptions qui ont été gravées par les Phéniciens tant dans leurs villes de la côte de Syrie-Palestine (Byblos ou Gebal, Sidon, Tyr, etc.) que dans leurs colonies et comptoirs plus ou moins lointains, notamment dans les îles (Chypre, etc.). On pense que les plus anciennes qui aient été retrouvées datent de la fin du IXᵉ s. av. J.-C. ; la plupart sont attribuées au Vᵉ s. et siècles suivants.

Ces inscriptions, généralement courtes, sont peu variées dans leur contenu ; les voyelles n'y sont pas notées ; aussi le phénicien est-il mal connu. Autant qu'on peut en juger en utilisant, avec les inscriptions, quelques noms propres notés dans les langues étrangères et le texte de Plaute (voir ci-après), il était distinct, mais peu différent du moabite et de l'hébreu.

Le phénicien des côtes de Syrie-Palestine, en pleine décadence dès la fin du IIIe s., semble avoir été complètement recouvert par l'araméen dans le dernier siècle avant l'ère chrétienne.

La plus importante des colonies phéniciennes a été Carthage ; le langage phénicien parlé par les Carthaginois s'était répandu dans le domaine qu'ils gouvernaient ; on le nomme au moyen du terme « punique », forme latine du nom « phénicien ». La puissance effective de Carthage se situe au plus tôt au VIe s. av. J.-C., mais les établissements phéniciens sur le même emplacement remontent à un temps beaucoup plus reculé. On date seulement du IVe s. les plus anciennes inscriptions puniques. L'écriture en est légèrement différente du phénicien continental ; la langue ne paraît pas représenter à l'origine un dialecte distinct. Sa prononciation est connue, très imparfaitement, par un petit texte continu : dix lignes déguisées en vers latins dans la pièce de Plaute « le Carthaginois » (*Poenulus*), à la fin du IIIe s. av. J.-C. ; la transcription et la coupe des mots dans les manuscrits causent des difficultés de lecture, mais la traduction latine rend le sens assuré. Les quelques lignes qui se trouvent dans les manuscrits entre le texte punique et la traduction latine paraissent être le même texte punique plus altéré et non, comme certains l'ont supposé, une version libyque.

A partir du IIe s. av. J.-C., sous la domination romaine après la chute de Carthage en 146, les inscriptions montrent une évolution de l'écriture et aussi de la langue (confusion des laryngales) ; on les appelle pour cette période néo-puniques. Certaines sont accompagnées de leur traduction en latin ; les dernières sont postérieures à la christianisation.

Le punique s'est conservé au moins jusqu'au IVe s. ap. J.-C. ; il est possible qu'il ait vécu encore plus tard, jusqu'à l'arabisation de son domaine (à partir du VIIe s.).

Hébreu. — L'hébreu est surtout pour nous une langue littéraire, celle de la Bible ; son importance pour la civilisation occidentale déborde beaucoup dans le temps et dans l'espace le cadre de la Palestine (environ 30.000 kilomètres carrés) dont il a été la langue parlée pendant environ mille ans. Il est bien connu, sous la réserve que de nombreux détails de la prononciation ne sont pas attestés à l'époque de son existence comme langue parlée (voir ci-dessous Écriture).

Le nom « hébreu » (racine *ʿbr* de l'hébreu) est une désignation de la population dont faisaient partie les fils d'Israël qui se trouvent établis dans le pays de Canaan, peut-être depuis peu de temps, vers la fin du XIIIe siècle av. J.-C. Dans la Bible, la langue des Israélites établis en Palestine (il n'est pas possible de savoir s'ils parlaient la même langue avant cet établissement) est nommée en de rares passages « (langue) juive » *yŏhūdīt*. Le mot « hébreu » n'apparaît, en hébreu tardif, comme désignation de la langue qu'à la fin du IIe s. av. J.-C. ; il est employé d'autre part en grec (*hebraïsti* « en hébreu ») ; il ne s'est pas implanté en hébreu postérieur, où le terme employé est « langue sacrée » *lŏšōn haqqōdęš*.

L' « hébreu ancien » ou « hébreu biblique » est la langue de la Bible et de quelques rares et courtes inscriptions (surtout : calendrier de Gezer en Judée, vers le IXe s. ; inscription du tunnel de Siloé à Jérusalem, vers la fin du VIIIe s. av. J.-C.). Son plus ancien monument, d'après les exégètes modernes de l'ancien Testament, est le chant de Debora (Livre des Juges, chapitre V), considéré comme antérieur à l'an 1000. Une grande partie de la Bible a été composée de la moitié environ du IXe s. à la fin du VIIe s. av. J.-C., c'est-à-dire après la séparation de l'État de David et de Salomon en deux royaumes (Israël au Nord, Juda avec Jérusalem au Sud) : poésies, paraboles, chroniques, légendes sur l'origine du monde et sur celle du peuple juif, prophéties, fragments de codes se sont peu à peu rassemblés. La destruction de Jérusalem et le départ des Juifs pour la captivité en Babylonie au début du VIe s. ont marqué le déclin de l'hébreu comme langue parlée de la Palestine ; par contre, ces circonstances

semblent avoir été plutôt favorables à la mise au point du recueil biblique. Celui-ci a pris forme après le retour partiel des tribus de Juda et de Benjamin de Babylonie en Judée (538 av. J.-C.), la reconstruction du temple de Jérusalem et l'établissement d'une solide théocratie (activité politique et intellectuelle d'Esdras, milieu du v[e] s.). Les vieux textes ont été à cette époque mis en ordre, amalgamés, complétés (en particulier les règlements religieux et juridiques), de manière à constituer la partie essentielle de la Bible, le Pentateuque. Mais déjà pendant cette période les derniers textes rédigés ne sont pas sans traces d'influence araméenne.

Lorsqu'un siècle plus tard, en 332, Alexandre conquiert la Palestine, l'hébreu paraît être à peu près exclu de l'usage parlé. Les juifs de Palestine parlent araméen, ainsi qu'une partie des émigrés (surtout en Babylonie). Ceux qui sont émigrés dans le monde hellénisé parlent grec; à partir du III[e] s. la Bible est traduite en grec par les juifs alexandrins (version des Septante) et c'est le grec qui sera la langue écrite des premiers chrétiens. Mais les milieux lettrés juifs ont continué à écrire l'hébreu de manière puriste jusqu'aux environs de l'an 100 av. J.-C.

Les juifs orthodoxes ne s'en sont pas tenus à la Bible. Les études considérables de générations nombreuses de docteurs (rabbins, de *rabbī* « monsieur ») ont abouti à la constitution d'un nouveau code juridico-religieux, la Michna (*mišnā*) « doctrine » ; soixante-trois petits traités y rassemblent des traditions historiques et des rites et coutumes qui s'étaient transmis oralement. La rédaction, confiée d'abord à la seule mémoire des docteurs, a ensuite été couchée par écrit en hébreu pendant les deux premiers siècles de l'ère chrétienne. Cet hébreu postbiblique de la Michna et de quelques autres textes est appelé aussi néo-hébreu, hébreu rabbinique ou « talmudique » (seule cette dernière dénomination est sans ambiguïté). Par certains traits il semble plus proche de l'hébreu parlé au moment de son extinction que les textes puristes contemporains de cette extinction, mais il paraît d'autre part influencé par l'araméen.

La Michna est entourée de volumineux commentaires en ara-

méen, avec quelques enclaves hébraïques, qui constituent avec elle la masse du Talmud.

Au moyen âge et jusqu'à nos jours, la langue savante des rabbins de tous pays dont la langue maternelle était l'arabe, l'espagnol, le français, l'allemand, etc., a été un hébreu parfois mélangé de formes araméennes, qui n'a plus rien d'une langue vivante, mais ressemble par son usage au latin scolastique ; certains auteurs réservent à cette langue savante le nom de néo-hébreu ou hébreu rabbinique ; d'autres lui appliquent ces noms, en même temps qu'à l'hébreu de la Michna, sans distinguer les différents moments postbibliques.

Vers la fin du XIXe siècle, le mouvement sioniste a tenté de ressusciter l'hébreu comme langue vivante, en généralisant l'emploi scolastique qu'en ont toujours fait les rabbins, pour écrire, enseigner, converser avec leurs confrères dans leurs déplacements en pays étrangers. L'hébreu a donc été passablement étudié, lu et écrit dans certains milieux, surtout en Europe orientale et en Palestine, mais il y a été et il y est très peu parlé.

Des mots hébreux ont pénétré en grand nombre dans les langages juifs tels que le yidich ; quelquefois ils forment les éléments de vocabulaires argotiques.

Écriture ; texte massorétique. — Les textes hébreux ont été notés depuis longtemps (vraisemblablement depuis l'époque d'Esdras) non en écriture cananéenne, mais dans une écriture araméenne qui est connue sous le nom de « hébreu carré ». L'écriture ainsi dénommée s'oppose aux écritures plus cursives, mais également à caractères séparés, qui sont dites « rabbiniques » ; la plus connue est celle à laquelle on a attaché le nom de Rachi, c'est-à-dire Rabbi Chelomoh Yishaqi (né à Troyes en 1040). L'hébreu a été quelquefois noté en caractères étrangers, tels que : arabes, grecs. Il est plus habituel que, inversement, des cursives hébraïques servent chez les juifs de différents pays à la notation de leurs langues maternelles. Dans ces cursives, pour les langues autres que l'arabe, les lettres notant des consonnes laryngales propres au sémitique ont été affectées avec les semi-voyelles à la notation des voyelles.

Cet usage moderne de notation des voyelles n'est qu'une extension de l'habitude qui s'est peu à peu implantée, à l'époque même où l'hébreu était parlé, de noter les voyelles longues au moyen de consonnes non prononcées (*matres lectionis* « mères de lecture »), ainsi *ʿyr* = *ʿīr* « ville », *ḥkmh* = *ḥọxmā* « sagesse ». L'ancienneté de cette notation est prouvée par sa présence sur d'anciennes inscriptions. Elle se généralise dans les textes de plus en plus récents (si bien qu'en néo-hébreu il en est fait quelquefois usage même pour des voyelles brèves). Toutefois, comme on ne connaît encore en fait de manuscrit ancien qu'un très court fragment (papyrus Nash, du IIᵉ siècle ap. J.-C.) et comme la tradition manuscrite ne commence authentiquement pour l'ensemble du texte biblique qu'au début du Xᵉ s. ap. J.-C., on ne peut jamais être assuré dans un cas particulier de l'ancienneté d'une *scriptio plena* « écriture pleine » avec consonne notant une voyelle.

La notation des voyelles par les matres lectionis est restée presque seule employée jusqu'à nos jours, en dehors du texte de la Bible.

La Bible au contraire est généralement pourvue (dans les exemplaires qui ne sont pas destinés à être lus solennellement à la Synagogue) non seulement d'une notation des voyelles, mais de tout un appareil destiné à assurer une lecture correcte et la connaissance de la cantillation pour l'usage liturgique. Le texte ainsi équipé est dit « massorétique » (de *māssōrẹt* « massora, tradition »). Le travail des massorètes, commencé au VIᵉ ou VIIᵉ s. ap. J.-C. n'a abouti qu'au XIIᵉ s. à l'adoption d'un texte unique chez tous les juifs. En fait nous possédons une notation de différents traits de l'hébreu prononcé (géminations de consonnes, affaiblissement des consonnes entre deux voyelles, qualité et quantité des voyelles, présence ou absence de la voyelle neutre *ö*, accentuation, etc.) mise par écrit mille ans après la constitution du canon biblique. Même en attribuant à la tradition liturgique une force parfaite de conservation, qui n'est pas vraisemblable, on n'atteindrait encore tout au plus que l'état de l'hébreu à la veille de sa disparition, au Vᵉ-IVᵉ s. av. J.-C. Or certaines nota-

tions des Septante au IIIe siècle, dans leur transcription grecque, montrent des discordances avec la Massora qui suffisent à empêcher qu'on prenne celle-ci comme témoin sûr pour cette époque. Le texte massorétique, si précieux qu'il soit, n'est donc pas une notation de l'hébreu ancien à laquelle on puisse se fier.

Le plus connu des systèmes de notation vocalique, et le seul qui soit encore utilisé presque partout à l'époque actuelle, est le système tibérien, ainsi nommé d'après l'école des savants de Tibériade en Palestine. Les voyelles y sont notées par des signes généralement placés en dessous des lettres. Ces signes ont sans doute été inspirés en partie par la vocalisation nestorienne du syriaque. — La notation tibérienne est prononcée de manière légèrement différente par les juifs méditerranéens (dits *sefardi* ou de rite portugais) et les autres (dits *aškenazi* ou de rite allemand).

Il existe aussi un système babylonien qui n'a pas survécu aux grandes écoles juives de Babylonie (environ IXe s.) ; il consiste en signes au-dessus des lettres. Enfin un autre système, écrit également au-dessus des lettres, a reçu le nom de « palestinien ».

Dans l'exemple suivant (Genèse, 42, 1) les parenthèses séparent les indications massorétiques, excepté celles qui concernent l'affaiblissement des consonnes, voir Observation 2.

w(ay)y(a)r(ȫ)' *y(ȧ)'(ă)q(ȱ)ƀ(,)* *k(i)y* *y(ẹ)ṡ(-)* *ṡ(ẹ́)ƀ(ẹ)r* *ƀ(ö)-*
et vit[1] Jacob[2] que il y avait du grain en
m(i)ṣr(ȧ)y(i)m (;) *w(ay)y(ȱ)'m(ẹ)r* *y(ȧ)'(ă)q(ȱ)ƀ* *l(ö)ƀ(å)n(ǻ)yw(,)*
Égypte, et dit Jacob à ses fils
l(åm)m(ǻ)h *t(i)ṗr(å)'w(ů)(.)*
pourquoi vous regardez-vous les uns les autres ?

Observations. — 1. En hébreu l'imparfait précédé de *wa* « et » (et ayant la consonne désinentielle géminée) exprime l'accompli (comme le parfait).

2. Dans le texte massorétique les consonnes occlusives non emphatiques sont généralement prononcées spirantes quand elles ne portent pas un signe de renforcement (lequel est le même que le signe habituel de gémination).

Araméen.

Le nom d'araméen (hébreu *'arāmīṯ* « en araméen ») s'applique à un ensemble de dialectes très proches entre eux dont certains ont eu un rôle littéraire important ; il n'y a pas de langue commune attestée, de même qu'il n'y a pas eu de grand État araméen dans l'histoire de l'antiquité. Dans la Bible et dans les documents accadiens les Araméens sont des tribus nomades dont l'aire de parcours paraît avoir été le Nord de l'Arabie jusqu'aux confins de la Syrie-Palestine et de la Babylonie. La plupart semblent s'être peu à peu établis sur les confins et jusque dans l'intérieur des régions civilisées (Pétra, Palmyre, Édesse, Damas, etc.), en se mélangeant avec des populations antérieurement établies.

Les citadins de langue araméenne, avant de fournir des théologiens au christianisme, ont donné des fonctionnaires aux Accadiens, et ils ont rempli les cadres administratifs de l'empire perse : celui-ci s'est servi de l'araméen comme langue de gouvernement dans les pays enlevés à la domination des Assyriens et des Babyloniens.

L'araméen s'est finalement substitué peu à peu à toutes les langues sémitiques du Nord : accadien, phénicien, hébreu. Sa vitalité et son usage littéraire ont limité l'extension du grec en Asie ; cependant il a été vaincu par le grec sur la côte méditerranéenne. L'époque de sa plus grande extension (environ 600.000 kilomètres carrés) se situe entre 300 av. J.-C. et 650 ap. J.-C. Il a été ensuite recouvert à son tour par une autre langue sémitique, l'arabe, et il n'est plus parlé de nos jours que par 200.000 individus environ.

L'araméen est très proche du cananéen. Il se divise en deux groupes de parlers : araméen occidental et araméen oriental. L'Euphrate et le désert de Syrie les séparent, des faits linguistiques les distinguent (ainsi, à l'Est, à partir d'une certaine époque, l'abandon du préfixe *y-* de la 3[e] personne à l'imparfait du verbe). Il y a donc lieu d'écarter l'ancienne division qui classait le chal-

déen comme dialecte oriental et le syriaque comme dialecte occidental. Les parlers de chaque groupe, souvent très proches entre eux par la grammaire et le vocabulaire, méritent des noms distincts à cause des différences de temps, de lieu, de religion, de civilisation ; la diversité des écritures répond à cette multiplicité des développements autonomes.

Araméen occidental ancien. — Ce nom collectif est donné provisoirement aux langues de quelques inscriptions dues à divers principicules de Syrie ; elles ne sont pas toutes semblables et peut-être pas toutes vraiment araméennes (ceci pour certaines inscriptions de *Zindẓirli*).

La plus ancienne inscription est celle d'un roi de *Ḥamā(t)* à mi-chemin entre Damas et Alep, datant du début du VIIIe s. av. J.-C. environ ; les inscriptions des rois de *Sam'al*, trouvées à *Zindẓirli* (Singerli), au nord d'Alep, sont un peu postérieures ; d'autres sont plus récentes encore.

L'écriture de l'araméen ancien (à 22 caractères) est distincte, quoique très proche, de l'écriture cananéenne ; on y a décelé l'amorce des modifications qui sont le trait commun des diverses écritures araméennes.

Araméen d'Égypte et araméen biblique. — On a retrouvé en Égypte divers documents araméens du Ve et du IVe s. av. J.-C. Les papyrus et pièces de poterie (ostraka) d'Éléphantine, qui contiennent toute une série de lettres et de contrats, ont permis de reconstituer l'existence d'une petite colonie juive parlant araméen, établie en ce lieu au Ve s. av. J.-C. La langue est la même que celle des parties araméennes de la Bible. L'écriture est à peu de chose près l'hébreu carré ; aucune écriture sémitique à l'encre n'est connue à une époque aussi ancienne.

Les textes araméens de la Bible sont : une partie du livre d'Esdras (vers 300 av. J.-C.) et une partie du livre de Daniel (premières décades du IIe s. av. J.-C.) ; pour être complet, il faut ajouter deux mots dans la Genèse donnés expressément comme araméens et un verset de Jérémie.

Au début du livre de Daniel il est fait mention de la langue

des Chaldéens (en grec *dialektos khaldaïkê*) parlée à la cour de Nabuchodonosor. Les commentateurs anciens, ignorant l'existence même de l'accadien (disparu de leur temps), attentifs d'autre part à la coïncidence entre le retour de la captivité de Babylone et le remplacement de l'hébreu parlé par l'araméen, ont confondu « langue des Chaldéens » et « araméen », de sorte que jusqu'à nos jours le nom de chaldéen, chaldéen biblique ou chaldaïque a été généralement appliqué à l'araméen biblique. (Le même nom de chaldéen a reçu encore d'autres usages, voir plus bas.)

Araméen palestinien. — Dans l'histoire postérieure de l'araméen palestinien (quelquefois nommé syrochaldaïque) il faut distinguer les documents juifs, samaritains, chrétiens.

Les documents juifs sont considérables; ils attestent le travail intense des rabbins pour vulgariser en langue parlée la Bible, par des traductions ou paraphrases dites « targoum » (*targūm* « traduction ») et la Michna, par des commentaires dits « guemara » (*gömårâ* « complément »). On appelle *talmūd* « instruction » l'ensemble de la Michna et d'une Guemara. Cette littérature targoumique et talmudique se divise en documents judéens (surtout des targoum faits dans la région de Jérusalem, antérieurs dans l'ensemble à la guerre d'Hadrien, 138 ap. J.-C.) et en documents galiléens (principalement la partie araméenne du talmud de Jérusalem, composée dans les écoles de Tibériade au IVe-Ve s. ap. J.-C.). L'araméen judéo-palestinien est quelquefois abusivement nommé chaldéen.

Les Samaritains, vestige de l'ancien royaume d'Israël, indépendants à l'égard de la théocratie de Jérusalem, ont conservé comme seul livre saint, jusqu'à nos jours (communauté de Naplouse), un texte du Pentateuque hébreu. Il est écrit dans le caractère samaritain qui dérive directement de l'ancien cananéen (avec emploi de matres lectionis). C'est également l'hébreu qui a surtout servi de langue savante aux docteurs samaritains. Cependant un targoum a été rédigé en araméen de la région de Samarie, et noté en écriture samaritaine ; il paraît dater du IVe s. ap. J.-C.; quelques autres textes savants ont été écrits dans le même araméen samaritain.

Les documents chrétiens galiléens se réduisent à quelques mots insérés en transcription dans l'Évangile grec.

Au ve-vi^e^ s. des chrétiens melkites (ou catholiques), opposés aux Syriens monophysites, ont fait des traductions de l'Évangile et de la Bible en dialecte judéen. Les manuscrits qui les contiennent ont été écrits surtout du viiie au xie s., mais il en existe du xve s. ; ils montrent une écriture syriaque d'un type spécial.

Au ixe s. l'araméen palestinien était probablement déjà éliminé par l'arabe de l'usage parlé.

Nabatéen. — Les Nabatéens (Nabathéens) de Petra (Arabie Pétrée) et de Bos(t)ra (dans le Hauran, à l'Est de la Palestine) ont eu du iiie s. av. J.-C. jusqu'à 106 ap. J.-C. un rôle important comme maîtres du transit entre l'Arabie et l'Occident. En 106 leur territoire devient la province romaine d'Arabie. On ne sait pas encore au juste si c'étaient en majorité des Araméens, ou des Arabes écrivant en araméen.

Le dialecte des inscriptions nabatéennes est un araméen proche du palestinien. Ces inscriptions, surtout votives et tombales, sont nombreuses à Pétra, à Bostra et dans les oasis septentrionales de l'Arabie *Tayma'* et *el-Ḥiğr* (*Madāin Ṣaliḥ*) ; il s'en trouve aussi en Phénicie, en Égypte, et même en Italie. On les date du début du i^er^ s. avant l'ère chrétienne au début du ive s. ap. J.-C. L'écriture nabatéenne est la première écriture sémitique sur pierre où les lettres soient liées en partie.

Des inscriptions nabatéennes mal tracées se rencontrent en grand nombre sur les rocs du Sinaï : tous ceux de ces graffiti sinaïtiques qu'on peut dater sont compris entre le milieu du iie et le milieu du iiie s. ap. J.-C.

Palmyrénien. — La ville de Palmyre (160 km. environ au N. E. de Damas) a été la capitale de caravaniers qui commandaient la route du Golfe persique à la Méditerranée : leur période prospère a commencé au i^er^ s. av. J.-C. ; Palmyre a été détruite en 273 ap. J.-C. par les légions romaines ; à ce moment il semble que son aristocratie était arabe ou arabisée.

Les inscriptions palmyréniennes sont nombreuses. La plupart ont été trouvées à Palmyre. La plus importante est celle qui con-

tient le tarif de l'octroi de cette ville. Ainsi que de nombreuses autres, elle est pourvue d'une traduction en grec. Des soldats aventuriers ont été écrire du palmyrénien jusqu'au bout de l'Europe (Angleterre). La date la plus ancienne qu'on relève dans les inscriptions est 9 av. J.-C.; la grande masse est de 128 à 271 ap. J.-C.

La langue est sensiblement pareille à l'araméen palestinien. L'écriture est assez proche de l'araméen d'Égypte; les lettres sont presque toujours séparées ; à partir du IIe s. un point au-dessus distingue le *r* du *d* qui était identique par le tracé.

Néo-araméen occidental. — La région montagneuse du Liban et de l'Antiliban n'a donné aucun texte littéraire, mais l'araméen y a mieux qu'ailleurs résisté à l'arabe. Des villages libanais parlaient encore araméen au XVIe-XVIIe s. ap. J.-C.; les parlers arabes de cette région conservent des éléments araméens. Un village chrétien de l'Antiliban, *Ma'lūla* (35 km. environ au N. de Damas) et deux villages musulmans voisins (1.500 âmes environ au total) usent encore d'un parler néo-araméen occidental (on l'a quelquefois nommé « syriaque occidental »).

Araméen oriental ancien. — On a des traces d'emploi de l'écriture et de la langue araméennes dès le IXe s. av. J.-C. en domaine accadien. Au VIIe s. les courts documents (surtout des titres araméens de tablettes dont le texte est écrit en accadien) sont nombreux ; une inscription suivie a été retrouvée à Assur. A l'époque perse l'araméen, se répandant, est attestée en Babylonie par des textes sur tablettes (Ve s.). Il se retrouve aussi dans des domaines excentriques : en Cappadoce, dans la haute vallée de l'Indus.

Ces documents ont une écriture de type ancien, à caractères séparés. La même écriture peu modifiée se lit encore sur de petites inscriptions retrouvées à Assur et à Hatra (Mésopotamie) qui paraissent être du IIIe s. ap. J.-C.

Syriaque. — Le parler de la ville d'Édesse (moderne Ourfa) en Mésopotamie septentrionale, a été une langue littéraire importante. Édesse, vieux centre politique (capitale de l'État indépen-

dant d'Osrhoène (fin du IIe s. av. J.-C. à milieu du IIIe s. ap. J.-C.), est devenu après sa conversion au christianisme, au IIe siècle, la métropole intellectuelle de l'Orient chrétien.

Des bords de la Méditerranée aux montagnes de la Perse la seule langue littéraire chrétienne a été le syriaque édessénien. Le nom « syriaque » est pris au grec ; les termes indigènes étaient les équivalents syriaques des mots : araméen, mésopotamien, édessénien ; le nom de syrien ne peut être appliqué qu'à l'arabe moderne de Syrie.

Quelques inscriptions funéraires représentent l'édessénien de l'époque païenne. D'autres inscriptions datent de l'époque chrétienne.

La littérature chrétienne d'Édesse a fleuri surtout du IIIe au VIIe siècle. Le monument le plus considérable en est la traduction complète de la Bible et des Évangiles, dite version « simple », en syriaque *pōšiṯṯa* d'après la prononciation dite orientale, *pōšiṯṯo* d'après la prononciation dite occidentale.

Du VIIIe au XIIIe s. la langue, déjà plus ou moins immobilisée auparavant par son usage littéraire, n'est plus qu'une langue savante : les chrétiens d'Orient adoptent l'arabe dans l'usage parlé et même les auteurs syriaques de la fin du Xe au XIIIe s. écrivent en arabe en même temps qu'en araméen.

Après le XIIIe s. le syriaque ne donne plus d'œuvres littéraires ; mais il reste la langue du culte chrétien et la pratique s'en est perpétuée dans le clergé jusqu'à nos jours.

Les circonstances religieuses qui ont régi le syriaque ont fait dès l'origine incorporer à son vocabulaire une grande quantité de termes grecs. Au Ve s. les querelles confessionnelles ont causé un schisme linguistique : tandis que les Melkites renonçaient au syriaque (voir page 109), à Édesse même il y avait séparation entre les Jacobites monophysites, qui se rattachaient à l'empire romain, et les Nestoriens diphysites, qui étaient surtout ressortissants perses. L'aspect linguistique de la querelle a été l'adoption par les deux partis d'habitudes de prononciation et d'écritures différentes. L'aspect politique en a été l'expulsion des Nestoriens d'Édesse à la fin du Ve s.

De leur nouveau centre, Nisibe, ils ont projeté par leur propagande religieuse la langue et l'écriture syriaques de la Perse à la Chine.

Écriture syriaque. — Dérivée de l'araméen ancien, connue d'abord par des monnaies et des inscriptions du Ier s. puis par des manuscrits du Ve s., elle a un caractère original de cursive, où la plupart des lettres sont liées à l'intérieur des mots ; un point diacritique est utilisé pour distinguer *r* de *d*. Le syriaque se lit de droite à gauche, mais quelques inscriptions ont aidé à démontrer que le tracé était fait de haut en bas, au moins à certaines époques.

L'écriture syriaque telle qu'elle a servi sans concurrente jusqu'à la séparation des Jacobites et des Nestoriens est appelée estranguela (estranghela), ou estranguelo (estranghelo). Les voyelles n'y ont jamais été notées que par des matres lectionis; des points distinguant certaines catégories de mots facilitent la lecture.

Les Nestoriens ont employé une forme peu modifiée de l'estranguela ; à partir du VIIIe s. ils ont adopté, surtout dans les textes liturgiques, un système de notation des voyelles par des points au-dessus et au-dessous de la ligne. Les Jacobites ont tracé une cursive plus grêle qu'ils ont munie (fin du VIIe ou VIIIe s.) de voyelles grecques placées sur ou sous la ligne.

L'écriture syriaque a eu une grande fortune, notamment en Extrême-Orient. On l'a appelée *karšūni* (karchounique) quand elle a servi, avec quelques additions, à la notation de l'arabe en pays araméen et à celle du langage des chrétiens de saint Thomas (côte du Malabar, Inde).

Harranien. — On n'a que quelques gloses du parler de la ville de Harran, au sud d'Édesse, restée païenne jusqu'au VIIIe s.

Talmud de Babylone. — Le principal dialecte araméen oriental, avec le syriaque, est la langue du Talmud de Babylone (*talmūd babli*) : celui-ci consiste en une guemara étendue qui a été rédigée dans les écoles juives de Babylone du IVe au VIe s. ap. J.-C. environ.

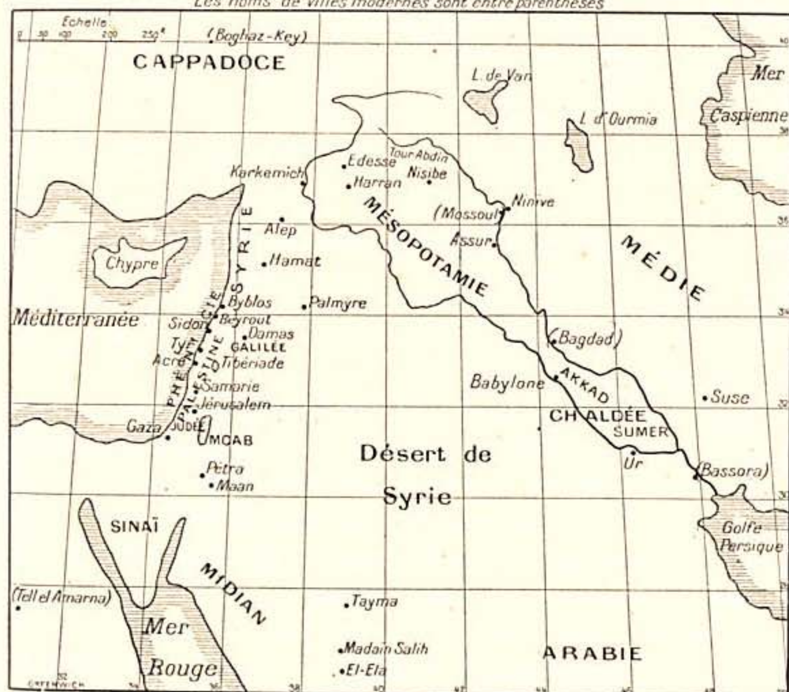
L'écriture est l'hébreu carré, sans autre vocalisation que les matres lectionis.

Mandéen. — Le domaine sud-est de l'araméen a vu se développer les littératures de diverses sectes qui ont mélangé des influences iraniennes aux influences judéo-chrétiennes. S'il ne reste rien des écrits araméens manichéens, les Mandéens (Mendéens, Mandaïtes, Mendaïtes) ou, par confusion avec une autre secte, Sabiens, Sabéens ont subsisté jusqu'à nos jours et ont préservé leur littérature écrite.

SÉMITIQUE ORIENTAL ET SEPTENTRIONAL

Pays et Villes dans l'Antiquité

Les noms de villes modernes sont entre parenthèses



La langue de cette littérature diffère peu de celle du Talmud de Babylone. Ses plus anciens monuments datent environ de la période VIIe-IXe s. ap. J.-C. L'écriture est une cursive araméenne à

caractères généralement liés, indépendante du syriaque ; l'usage des matres lectionis est étendu au point que toutes les voyelles sont notées.

Néo-araméen oriental. — Des parlers qui n'ont pas eu anciennement d'importance littéraire se sont conservés vivants jusqu'à nos jours dans la région montagneuse qui forme l'extrémité nord-est du domaine araméen, où ils s'entremêlent avec des parlers kurdes (iraniens). Ce néo-araméen oriental, souvent appelé néo-syriaque, ou syriaque vulgaire, est la langue de 200.000 individus environ. Les jacobites de la Mésopotamie septentrionale, dans le district de *Ṭūr ʿabdīn*, parlent le « torani » qu'on a appelé aussi le mésopotamien. Plus à l'Est, le néo-araméen est parlé surtout par des nestoriens, mais aussi par un petit nombre d'uniates (ralliés à la catholicité et qu'on appelle chaldéens), de jacobites et de juifs ; le principal dialecte est l'« ourmien », dont les missionnaires protestants et catholiques ont fait à la fin du XIXe s. une espèce de langue littéraire (avec écriture nestorienne).

SÉMITIQUE (OCCIDENTAL) MÉRIDIONAL

Le sémitique méridional est remarquable en phonétique par l'abondance des consonnes (en particulier existence des interdentales) et le passage de *p* à *f*.

En morphologie, les traits distinctifs principaux sont les suivants : il existe, comme en vieil accadien, une flexion nominale à trois cas ; le pluriel des noms est souvent exprimé au moyen de collectifs, formés surtout par modifications internes des radicaux singuliers (pluriels internes ou brisés) ; dans le verbe, on distingue généralement un mode subjonctif.

Le domaine ancien du sémitique méridional est l'Arabie, avec la partie de l'Afrique qui a été colonisée par des habitants de l'Arabie. En Arabie même il faut séparer la région des langues sud-arabiques et la région de l'arabe ; dans celui-ci il faut encore distinguer de l'arabe proprement dit (dont le domaine ancien est l'Arabie centrale) les dialectes attestés par des inscriptions dans le Nord de l'Arabie et sur ses confins.

ARABE.

Inscriptions du Nord de l'Arabie. — Ces inscriptions se rencontrent au nord du Hedjaz, dans les oasis de *Taymā'* (Teima), *el-Ḥiğr* (*Madāin Ṣāliḥ*), *el-'Olā* (el-'Ela, el-'Ula, Dedan).

Certaines de ces inscriptions contiennent la mention des rois de *Liḥyān* ; on les situe à une époque allant du IIe ou Ier s. avant au IVe-VIe après J.-C. La langue lihyanite est proche de l'arabe. L'écriture est de type sudarabique ; la direction est horizontale, de droite à gauche.

Dans les mêmes endroits et aussi plus au Nord se relèvent des graffiti en une écriture moins soignée, tracés en différentes directions (souvent de haut en bas) ; on leur attache le nom de thamoudéens (tamoudéen, thamoudique, thamoudite) d'après le nom des Thamoud (*þamūd*) que le Coran situe dans cette région et dont les *Liḥyān* étaient peut-être une fraction. On a aussi appelé ces graffiti protoarabes, protoarabiques.

D'autre part, au Nord du domaine nabatéen (au S.-E. de Damas), les habitants du *Ṣafā* ont tracé des graffiti dans leur langage arabe septentrional (à côté de graffiti en grec), probablement pendant les trois premiers siècles de l'ère chrétienne. L'écriture safaïte (safaïtique) est analogue à l'écriture lihyanite ; la direction n'en est pas fixe.

Arabe ancien.— La langue arabe proprement dite, et dans cette langue le nom même des Arabes (racine *'r b*), apparaissent d'abord dans une inscription funéraire royale à En-Nemara (partie sud du Safa) ; elle est datée de 328 ap. J.-C., écrite en caractères nabatéens. L'écriture arabe, encore dénuée de tout point diacritique, apparaît dans l'inscription de *Zabad* (S.-E. d'Alep), datée de 512 et dans celle de *Ḥarrān*, au Sud de Damas, datée de 568.

A peu près à la même époque une langue littéraire (mais non encore écrite), celle de la poésie antéislamique, était florissante dans l'Arabie centrale. Une espèce de compromis entre la poésie antéislamique et le parler de la population des *Qoraïš* de la Mecque a été la langue du Coran (*qur'ān*), qui est l'ensemble de la prédication de Mahomet, dans la première partie du VIIe siècle.

L'arabe littéraire (ancien, classique, littéral, savant, coranique) est une des langues les plus importantes que connaisse l'histoire. D'innombrables auteurs ont écrit la langue des poèmes antéislamiques et du Coran : volumineux commentaires du Coran, recueils de traditions et de prescriptions, poésies de toutes sortes, ouvrages historiques, dictionnaires, traités de sciences exactes, contes, livres de voyages et d'aventures, tous les genres sont représentés. Aucun autre dialecte arabe n'est devenu langue littéraire ; tout au plus le vocabulaire s'est-il assez renouvelé après quelques siècles dans certaines régions pour qu'on puisse parler du « moyen arabe » ; c'est encore l'arabe classique, plus ou moins correctement lu et pourvu des néologismes nécessaires, qui est la langue de la presse contemporaine (voir p. 117). Le domaine de l'arabe littéraire ne se borne pas aux pays où l'arabe s'est répandu comme langue parlée ; le Coran ne devant, aux yeux des musulmans orthodoxes, être ni traduit ni transcrit, le Coran arabe a voyagé au loin avec l'Islam. L'arabe a été la langue écrite de peuples peu civilisés ; à des peuples musulmans qui ont écrit leur propre langue, il a fourni un très grand nombre de termes religieux et intellectuels (turc, persan, etc.).

L'arabe classique a tous les caractères d'une langue littéraire conservatrice ; la régularité schématique de l'ensemble grammatical (ainsi la notation de trois timbres vocaliques seulement, avec quantité brève ou longue) ; les règles minutieuses d'une syntaxe abondante en distinctions subtiles, presque toutes inexistantes en arabe parlé moderne ; la surabondance d'un lexique où se cumulent les vocabulaires particuliers des lieux et des époques et les trouvailles des auteurs stylistes, tout porte la marque d'une langue dès l'origine savante et non d'usage journalier.

Écriture. — L'écriture arabe, qui continue le nabatéen, mais a été influencée par le syriaque, s'est perfectionnée dès le VII[e] siècle en une cursive rapide, où la plupart des lettres sont jointes ; des points diacritiques ont distingué les caractères dont le tracé se confondait et les consonnes qui, existantes en arabe, sont inconnues du sémitique septentrional. Le total des caractères est de

vingt-huit. L'écriture ordinaire *nasḫ(i)* se distingue d'une forme raide, surtout monumentale, le coufique.

Les grammairiens exégètes, soucieux surtout de conserver rigoureusement exacte la prononciation du Coran, ont adopté des signes complémentaires, principalement un signe de gémination et des signes pour noter les voyelles ou l'absence de voyelle. Un système rappelant le syriaque nestorien, avec des points, n'a pas persisté. Celui qui a été adopté au milieu du VIII[e] s. environ emploie des lettres sur et sous les consonnes, comme le système jacobite, mais les emprunte à l'alphabet arabe (ainsi *w* pour *u*). Les exemplaires du Coran sont toujours voyellés. La vocalisation peut aussi servir à l'enseignement. Mais l'ensemble des manuscrits et imprimés arabes est écrit sans voyelles, souvent même sans signe de gémination. Les conséquences de cette situation sont multiples. D'abord un texte arabe ne peut être lu correctement et même bien compris que par un lettré. D'autre part, le consonantisme étant resté très ferme dans l'évolution de l'arabe parlé, tandis que les voyelles n'ont pas cessé de se réduire plus ou moins suivant les dialectes, un texte non voyellé peut à la rigueur se lire autrement qu'avec la prononciation classique ; un article de journal, un conte, et surtout une lettre peuvent être, quoique écrits en arabe littéraire, déchiffrés tout bas ou même tout haut en parler moderne. Cette altération conditionnelle de la langue savante, malaisément tangible, personnelle à chaque lecteur peu instruit, est finalement une des causes de la persistante juxtaposition de cette langue à la langue parlée.

L'écriture arabe a eu une vaste fortune, dépassant encore celle de la langue littéraire ; elle a servi et sert encore à noter des langues variées (turc, persan, hindoustani, etc.) ; l'espagnol qui a été écrit en caractères arabes par des semi-arabisés est dit « aljamia, aljamiado ».

D'autre part l'arabe a été quelquefois noté dans d'autres écritures : il est écrit en caractères hébraïques par les juifs, en caractères latins par les Maltais (sur le *karšūni*, voir p. 112).

Arabe moderne. — Des dialectes variés existaient au temps de

Mahomet, en Arabie centrale ; les armées conquérantes de l'Islam rassemblaient avec les Arabes du centre des Arabes du Sud et du Nord dont les parlers étaient différents. Toutefois ces armées ont dans l'ensemble transporté une langue sensiblement une ; dans l'état actuel des études on ne peut pas assigner de sources différentes aux parlers modernes ; tout semble s'être passé comme si un seul langage s'était répandu sur la partie du monde où l'Islam a déferlé en l'arabisant, et y avait évolué ensuite en parlers divers. L'Arabie elle-même a plus de chances d'avoir préservé des divisions dialectales anciennes ; mais elle est encore trop mal explorée pour qu'on puisse bien en juger.

L'arabe parlé (moderne, vulgaire) est de nos jours la langue de 30 millions d'individus environ. Son domaine, se restreignant sur certains points (Europe), s'étendant sur d'autres (Afrique), a été sensiblement constant depuis le XIVe s. (environ 14 millions de kilomètres carrés).

Sur ce domaine il y a de très nombreux parlers : chaque tribu de nomades (bédouins), chaque canton de campagnards sédentaires (*fellaḥ*, *faḥṣi*), chaque ville, et chaque élément citadin dans les villes complexes, a une manière spéciale de parler arabe. Le langage local s'écarte plus de l'arabe ancien dans les régions les plus éloignées de l'Arabie ; dans chaque région, les parlers de nomades et campagnards sont plus conservateurs ; les citadins innovent plus.

Malgré sa diversité, l'arabe parlé est resté assez un pour que les gens de langue arabe puissent arriver à communiquer entre eux au prix d'efforts d'accommodation peu considérables, de Zanzibar à Mogador. Au contraire de l'arabe littéraire, c'est une langue d'un maniement grammatical simple.

L'arabe parlé n'est presque nulle part connu par des témoignages anciens. L'étude se réduit presque partout à l'arabe moderne, observé depuis le XIXe s. par les savants européens.

Les principales divisions dialectales du domaine arabe répondent à des divisions géographiques naturelles et en partie à d'anciennes régions linguistiques et nationales distinctes, recouvertes à différentes époques.

1. Arabie (parlers arabiques).

Vers le Nord : les dialectes du Hedjaz, en particulier le parler de la Mecque (*Mäkka*), proche des parlers citadins de Syrie. Le désert de Syrie (ci-dessous 3) pourrait se rattacher à l'Arabie du Nord. A l'Est du Hedjaz, parlers du Neḍjd, mal connus.

Au Sud : dans le Yémen, l'arabe d'Aden, qui est aussi parlé en Afrique chez les Somali comme langue seconde. Plus à l'Est, les parlers du Haḍramaut et du *Daþīna* (datinois). Plus à l'Est encore, les parlers de l'Oman (Omanais) ; on a décrit surtout un parler du centre (à l'Ouest de Mascate) ; un parler de la corne nord s'appelle le *šaḥi* ; le parler de Zanzibar, en Afrique, est de l'omanais.

Pour cette région, il faut tenir compte de l'islamisation des îles Comores et, partiellement, de Madagascar, où un sabir à base d'arabe a été en usage.

2. Iraq (*'Irāq*), ancienne Babylonie (dialecte iraqois) ; en particulier parler de Bagdad. Mésopotamie : parlers de Mossoul, de Mardin. Argot *qurbāti* en Perse méridionale. Peut-être 3.000.000 d'habitants.

3. Syrie (arabe syrien, syrien) : parlers d'Alep, de Beyrouth, de Damas, des villages du Liban (libanais). Palestine (arabe palestinien) : Jérusalem, les campagnards. Le désert de Syrie : parlers bédouins. Environ 4.000.000 d'individus.

4. Égypte (arabe égyptien, égyptien) ; en particulier, parler du Caire (cairote), argot *ḥalebi* ; la haute Égypte est mal connue. Des évaluations relativement récentes donnaient 8 millions d'habitants à l'Égypte ; mais les dernières statistiques montent à 14 millions.

De l'Égypte dépend linguistiquement le Soudan, avec des populations mélangées de bédouins arabes envahisseurs et de chamites (Est), de nègres (Ouest) ou de négroïdes. Autour du lac Tchad, les Arabes sont appelés par les nègres « choa » (*šoa*, *šuwa*, etc.), nom souvent donné aux parlers arabes du Bornou et du territoire du Tchad ; le parler du Ouaday fait partie du même ensemble. Les parlers orientaux du Soudan (sur la rive gauche du Nil) sont encore très mal connus.

5. Maghrib (c'est-à-dire, en arabe, Occident) ; dans les anciens ouvrages, Barbarie.

Le grand domaine des parlers maghribins (magrébins, mog(h)-rébins) s'étend des confins égyptiens à l'Atlantique ; il se subdivise en groupes de parlers dont quelques-uns ont commencé à se développer déjà à partir du VIIᵉ s., mais dont l'ensemble s'est constitué surtout depuis le XIᵉ s.

Il faut prendre garde, dans la nomenclature, au cas où une ville donne son nom à une région ; ainsi « algérois » désigne les parlers de la ville d'Alger (Alger-musulman et Alger-juif) ou les parlers de la région d'Alger ; « algérien » dénomme l'arabe de toute l'Algérie).

a. Parlers des îles. A Malte, le « maltais », parlé par des chrétiens maltais, et écrit en caractères latins.

En Sicile aussi l'arabe a été parlé (domination arabe au Xᵉ siècle) ; des diplômes siciliens arabes ont été retrouvés et publiés.

L'arabe a laissé des emprunts nombreux dans les parlers romans à Pantellaria (où il a été en usage jusqu'au XVIIIᵉ s.), et aux îles Baléares.

b. Libyen (libyque) et tripolitain (en particulier, ville de Tripoli). Environ 1 million d'individus.

c. Tunisien, en particulier ville de Tunis. Environ 2.000.000 d'individus.

d. Algérien ; il faut distinguer, de l'Est à l'Ouest, les parlers des trois départements de Constantine, Alger, Oran ; d'autre part, les parlers des villes (dans certaines se rencontrent les Maures, population d'origine très mélangée). Il faut distinguer aussi du Nord au Sud diverses régions : le *sāḥel* (bande côtière), le *täll* (bordure nord des hauts plateaux) ; le Sahara. Environ 3.500.000 individus.

e. Andalou ou andalous (arabe d'Espagne). L'arabe a été parlé en Espagne méridionale à partir du VIIIᵉ s. jusqu'au XVIᵉ s. L'arabe d'Andalousie est connu par des poèmes en langue vulgaire du XIIᵉ s., et par des études d'Espagnols chrétiens au XIIIᵉ s. et à la fin du XVᵉ s. De nombreux emprunts arabes sont restés en espagnol.

f. Marocain (arabe marocain). L'arabe est parlé au Maroc par deux cinquièmes des habitants (soit environ 2.400.000) surtout dans les régions basses.

g. Hassani ; ainsi nommé d'après la tribu conquérante des *Banī-Ḥassān* ; c'est l'arabe de la Mauritanie et des territoires à l'Est de la Mauritanie, jusqu'à Tombouctou ; cet arabe élimine de plus en plus les parlers berbères et soudanais au nord du fleuve Sénégal, et sert de langue de relation à de nombreux musulmans.

SUDARABIQUE.

Sudarabique ancien (appelé aussi himyarite, sabéen, yéménite, etc.). Cette langue est représentée par des inscriptions nombreuses, surtout dans le Yémen et dans les oasis au Nord du Hedjaz.

L'étude de ces inscriptions permet de reconstituer en partie l'histoire ancienne de l'Arabie occidentale.

Un État florissant a été constitué dans le Sud-Ouest de l'Arabie par les Minéens (*Ma'in*) avant tous autres états connus en Arabie, mais à une époque mal déterminée ; les vraisemblances paraissent être pour le VIIIᵉ s. av. J.-C. Les Minéens, maîtres du commerce par terre de l'Océan indien vers le Nord par la côte arabique de la mer Rouge, avaient des colonies septentrionales, au contact du monde cananéo-araméen ; le minéen y a été écrit sur la pierre, ainsi qu'en Yémen. Lorsque la puissance a passé ensuite aux Sabéens (*Saba'*), qui semblent n'avoir pas eu de colonies au Nord, c'est le dialecte « sabéen » qui a prédominé dans les inscriptions yéménites. La période suivante, illustrée à son début par les luttes des Sabéens et des Himyarites (gens de *Ḥimyar*), a vu persister la prédominance linguistique du sabéen. Quand les Abyssins établissent au début du IVᵉ s. ap. J.-C., pour une courte période, leur prédominance sur le Yémen, il n'en résulte pas non plus de changement dans la langue des inscriptions.

Quelques inscriptions ne provenant pas de l'État principal attestent un dialecte occidental, celui du *Qaṭabān* (au Nord du site d'Aden) et un dialecte oriental, celui du Hadramaut (*Ḥaḍramawt*).

Les dernières inscriptions sudarabiques, au VIᵉ s., ont précédé de peu l'établissement de gouverneurs perses ; eux-mêmes ont été chassés au VIIᵉ siècle par la conquête musulmane.

La tradition de la langue écrite sudarabique, incomprise des Arabes et des Sudarabiques modernes, s'est éteinte peu à peu, sans qu'aucune littérature ait été conservée en dehors des inscriptions.

La langue de ces inscriptions, dans la mesure où une écriture sans voyelles permet de la connaître, apparaît proche de l'arabe, mais avec des particularités phonétiques et morphologiques nettes, et un lexique indépendant.

Écriture sudarabique. — Cette écriture monumentale est une ramification originale de l'alphabet sémitique, indépendante, semble-t-il, des écritures du Nord. Les rapports de dérivation entre elle ou son prototype et les écritures lihyanite et safaïte ne sont pas encore élucidés. Cette écriture est remarquable par les formes géométriques régulières de nombreux caractères. Le tracé est horizontal, la direction de droite à gauche prédomine, mais les inscriptions serpentines, droite à gauche, puis gauche à droite (tracé boustrophédon) sont fréquentes. Il y a vingt-neuf signes de consonnes différentes.

Sudarabique moderne. — Si l'arabe a recouvert la région des inscriptions sudarabiques jusqu'à présent connues, des parlers non arabes ont subsisté sur le domaine immédiatement contigu à l'Est. On distingue : le *mahri* (mehri), dans le Mahra, bande côtière à l'Est du Hadramaut (de 51° à 52°15′ Est de Greenwich); le *qarawi* ou *grawi* (*šḫawri*, *ḥakili*, *eḥkili*) contigu au mehri à l'Est, dans un petit district montagneux ; le *soqoṭri*, parler de l'île de *Soqoṭra* et des îles voisines.

D'après un renseignement isolé, un parler, celui des *Minhālī*, formerait une transition entre le mehri et le parler arabe du Hadramaut.

Langues éthiopiennes.

L'ensemble des langues éthiopiennes représente une avance du sémitique en Afrique, antérieure de beaucoup à l'expansion de l'arabe : invasion par des colons venus de l'Arabie du Sud, commencée sans doute plusieurs siècles avant l'ère chrétienne ; État constitué au Ier s. au plus tard. Les diverses langues du groupe

éthiopien semblent dues à l'évolution parallèle de plusieurs dialectes sudarabiques proches entre eux, soumis en domaine africain à l'influence des parlers chamitiques qu'ils ont plus ou moins lentement éliminés. Une classification rigoureuse n'en a pas encore été faite. (Voir la répartition sur la planche 4.)

Le nom « éthiopien » désignait en grec différents éléments africains ; les gens de l'Arabie du Sud installés en Afrique se sont appliqué ce terme grec. Le nom indigène de l'ensemble de la population immigrée est une forme de la racine *ḥbš*, d'où Abyssinie, Abyssin, qui sont les noms propres, sinon officiels, du pays et de ses habitants.

L'empire abyssin a, de nos jours, depuis les conquêtes de Ménélik, une superficie de un million de km. carrés environ, auxquels il faut joindre, pour envisager le domaine éthiopien complet, les 75.000 km. environ de la colonie italienne de l'Érythrée. La population est difficile à évaluer ; dix millions sont sans doute un minimum. Sur l'ensemble de l'empire, les langues sémitiques recouvrent à peu près le quart de la superficie, dans les régions où la population est de beaucoup la plus dense ; on peut leur attribuer environ cinq millions d'individus ; le reste parle chamitique et, pour une faible part, soudanais.

Ethiopien ancien (*guèze*). — L'éthiopien ancien (éthiopien classique, éthiopien) a pour nom indigène *gö'öz*, qu'on a diversement accommodé en français : guèze, ghèze, ghez, gheez, ge'ez, etc. ; au XVIe s. des savants européens l'avaient appelé chaldéen, chaldaïque sans aucune raison valable.

Le guèze apparaît d'abord sur des inscriptions qui, pour la plus grande partie, datent du IVe s. ap. J.-C., et se trouvent à Axoum, capitale de l'Abyssinie à cette époque. Elles sont dues presque toutes au roi Aeyzanas (*'Ezana*) qui a introduit le christianisme dans ses États. La Bible paraît avoir été traduite en éthiopien au Ve siècle.

L'empire d'Axoum s'est disloqué vers le Xe s. A partir de ce moment, tandis que la langue parlée va continuer à évoluer en dehors des influences savantes (voir ci-dessous *tigrigna*), la langue

écrite et savante se transporte avec la culture chrétienne vers le Sud ; après une période de puissance du Lasta (v. p. 146), un empire abyssin est reconstitué à la fin du XIIIe s. dans la province du Choa, avec l'amharique (v. ci-dessous) comme langue parlée. Les plus anciens manuscrits éthiopiens connus ne datent que de cette renaissance.

Le guèze a survécu comme langue liturgique et savante jusqu'à nos jours. Sa littérature religieuse et historique, ne cessant d'augmenter, a longtemps empêché le développement d'autres langues littéraires.

Le guèze est proche de l'arabe et du sudarabique ancien. Il se distingue en particulier par une certaine réduction du consonantisme.

La prononciation liturgique moderne conserve des géminations anciennes non notées par l'écriture, mais l'ensemble du consonantisme y est amharisé, c'est-à-dire privé surtout des consonnes laryngales. Il en est résulté de nombreuses confusions dans l'orthographe des scribes. Par là et par certains détails de syntaxe (dus en partie à l'influence de l'arabe, dont beaucoup de textes étaient traduits) le guèze littéraire a un caractère un peu artificiel.

Écriture. — La naissance de l'écriture éthiopienne, qui sert pour les langues sémitiques d'Abyssinie en général, semble devoir être datée du IIIe s. ap. J.-C. ; elle est donc antérieure à l'époque de *'Ezana* (IVe siècle), mais elle n'était pas encore uniquement employée au début de son règne ; en effet la principale des inscriptions de ce roi a, au revers d'un texte grec, la traduction éthiopienne, d'abord en caractères sudarabiques, puis en caractères éthiopiens non vocalisés. Ceux-ci sont presque semblables aux précédents, mais ont des formes plus arrondies qui font soupçonner qu'ils dérivent d'un tracé cursif. Une autre inscription du même roi est le plus ancien exemple de l'écriture éthiopienne vocalisée, qui par la suite est restée seule en usage.

Les consonnes nues se prononçant avec *ă*, six autres nuances vocaliques se marquent par des signes accrochés aux consonnes, ainsi *bū*, *bī*, *bā*, *bē*, *bö* (ou *b*) et *bō*, le signe étant sensiblement

constant pour chaque voyelle avec les différentes consonnes. L'alphabet-syllabaire a donc sept formes pour chacune des vingt-six consonnes. Il n'y a pas de signe de gémination. La direction de l'écriture est : de gauche à droite.

Tigrigna ou *tigray*. — Les deux noms signifient « tigréen », le premier en amharique, le second en tigrigna même, « Tigré » étant le nom indigène du domaine abyssin au Nord du fleuve Takkazé. Langue de la région dont Axoum est le centre, le tigrigna est à considérer comme du guèze évolué. Il diffère légèrement suivant les provinces (au Nord, dialecte du *Ḥamasen*). Très peu écrit jusqu'à nos jours, le tigrigna a été recueilli oralement dans la dernière partie du XIXe s. Comme l'administration italienne de l'Érythrée en fait usage, il tend à s'écrire davantage, et à s'étendre aux dépens des parlers voisins. Il est parlé par environ 500.000 individus (chrétiens).

Tigré. — Dans le Nord du domaine tigréen est parlé un idiome dans lequel le mot *tigre* désigne à la fois le pays et la langue elle-même ; ce terme, transporté dans les ouvrages européens, provoque malheureusement des confusions avec le tigrigna ; il faut bien prendre garde que tigré et tigrigna sont deux langues distinctes. Sur une partie seulement de son domaine le tigré a le nom de *ḫasa* (en arabe *ḫāsiya*, *ḫāsi*, d'où la transcription approchée « khasy » ; en bedja *to-hasa*, qu'on a quelquefois noté « tahase »).

Le tigré est le représentant d'un dialecte proche du guèze. Sa littérature orale a été recueillie récemment. Quoiqu'il ne soit pas langue écrite, il fait figure de langue de civilisation, gagnant sur des parlers voisins et servant de langue seconde à des Chamites et à des Soudanais. Il est la langue unique de 100.000 individus environ, consistant en tribus musulmanes qui habitent la région côtière de Massaoua à Souakin et les îles Dahlak.

Amharique. — C'est la langue parlée dans la majeure partie du haut-plateau abyssin, dans l'ensemble entre le Takkazé (sauf sur une petite longueur) au Nord, l'Abbay (Nil Bleu) et les abords

de l'Aouache (*Hawāš*) au Sud. Il reçoit son nom de la province centrale *Ambarā* (ancien *Amḫarā*), soit sous la forme indigène en -*gna* (d'où amharegna, amarigna, etc.), soit sous la forme latino-française en -*ique*.

L'amharique, indépendant du guèze dès l'origine, s'est écarté passablement du type sémitique ordinaire, sous des influences chamitiques ; en particulier il a adopté l'ordre des mots couchitiques (voir p. 142).

« Langue du roi » depuis la fin du XIIIe s., l'amharique semble avoir été employé depuis longtemps, comme il l'est de nos jours, dans la correspondance officielle. L'usage écrit qu'on en faisait a imposé l'adoption d'un signe spécial qui s'attache à sept signes de l'alphabet éthiopien pour noter des sons mouillés qui y sont fréquents, mais n'existaient pas en guèze.

Les plus anciennes poésies conservées par écrit datent du XIVe-XVe s. Au XIXe s. a commencé à se constituer une littérature amharique ; elle consiste surtout en traductions du guèze ; l'usage de cette langue savante est désormais menacée par l'habitude qui se répand d'écrire de plus en plus la langue parlée officielle.

L'amharique, langue de civilisation envahissante, a recouvert des parlers couchitiques divers ; il est généralement langue seconde pour les gens qui ont conservé un parler de ce groupe. Il se répand de plus en plus dans les régions méridionales de l'empire ; en particulier la capitale moderne, Addis-Ababa, elle-même bâtie dans une campagne galla, attire et amharise de nombreux individus. L'amharique a gardé en s'étendant une remarquable unité ; il a plutôt quelques particularités provinciales que de vrais dialectes ; les puristes préfèrent encore le parler de Gondar, ville située dans le Nord de la région amhara et qui a été capitale du XVIIe au XIXe s. ; on note comme particuliers les langages du Godjam (godjamite) et du Choa (choanais).

Gafat. — C'est le langage propre d'un petit district au S.-O. du Godjam, sur la rive droite de l'Abbay, dernier reste du domaine anciennement plus étendu de la population appelée Gafat. Signalé encore au XIXe s., ce langage a peut-être tout à fait

disparu depuis. Il est mal connu ; sa position parmi les langues éthiopiennes est mal déterminée.

Argobba. — Également très peu répandu et très mal connu, ce langage paraît devoir être situé dans deux districts du nom d'Argobba : l'un, dans les montagnes à l'Est du Choa, est habité par des musulmans ; l'autre, au Sud de Harar, semble avoir été peuplé par des émigrants de la première région.

Harari. — C'est la langue des citadins musulmans (20.000 environ) de la ville de Harar (*Adari* est le nom somali de la ville et du langage). Quelques textes, encore non traduits, sont attribués au XVIe s. (en écriture arabe) ; des textes très courts ont été notés au XIXe s. en écriture éthiopienne. L'étude de ce parler, qui est menacé par l'extension de l'amharique, est encore insuffisante.

Gouragué (gouraghé, guragie). — C'est un ensemble de dialectes variés parlés dans la région de ce nom (environ 10.000 km. carrés), par des éléments chrétiens, musulmans et païens. Le dialecte principal est le tchaha (*tšöha*) ; il n'est pas seulement le parler du canton de ce nom, mais il fait figure, pour tout un district, de langue littéraire dans la poésie orale.

Le gouragué représente l'avance la plus méridionale du sémitique abyssin, attestant une colonisation ancienne non datée exactement, que l'invasion galla du XVe-XVIe s. a interrompue pour quatre siècles. On n'en possède pas de documents antérieurs au XIXe siècle.

Égyptien.

La langue ancienne de l'Égypte a une histoire plus longue en un même lieu qu'aucune autre langue. Elle nous est connue historiquement peut-être depuis les abords de 4000 av. J.-C. La même

langue évoluée était encore vigoureuse au VIIe s. ap. J.-C. et se survit de nos jours comme langue liturgique.

L'Égypte est ainsi nommée d'après le grec *Aigüptos* ; le nom indigène ancien avait les consonnes *kmt*. Elle comprend, avec le Delta du Nil au Nord, une longue et étroite bande de terrain cultivable sur les deux rives du fleuve ; la limite sud de cette bande s'est déplacée à différentes époques suivant les degrés de puissance du gouvernement égyptien. La superficie cultivable totale est de 30.000 kilomètres carrés environ. On n'a aucune raison de croire que, malgré certaines expéditions lointaines vers le Sud ou dans le Nord (notamment, conquêtes en Asie antérieure), la langue ait jamais débordé ce domaine (sauf sans doute dans une partie de la presqu'île du Sinaï, où se trouvent de nombreuses inscriptions égyptiennes).

Cette région riche et civilisée a dû être peuplée aux périodes prospères de l'antiquité d'une manière sensiblement aussi dense que de nos jours (soit autour de 10 millions d'habitants environ, voir p. 119).

L'égyptien ancien se caractérise dans l'ensemble chamito-sémitique comme une langue depuis longtemps très évoluée et ayant suivi une voie à part dans le système du verbe. L'ancienne flexion verbale avec désinences préfixées y a disparu dès les plus anciens textes, sauf peut-être quelques traces, si toutefois elle y a jamais existé ; une conjugaison riche en nuances s'est constituée par la jonction de radicaux, nus ou augmentés de suffixes variés, avec des éléments pronominaux suffixés : *sḏm-k* « tu entends », *gm-tw-f* « il est trouvé ». — Les racines trilitères prédominent (voir p. 92). Il est probable que le consonantisme ancien a été très proche de celui du sémitique ; mais il s'est altéré dans la dernière période du développement de la langue.

Égyptien ancien.

L'histoire de l'égyptien est inséparable d'une histoire sommaire de l'Égypte et de l'écriture hiéroglyphique (en grec : « de gravure sacrée »).

Période primitive. — Pour l'histoire c'est la période dite prédynastique, où l'Égypte n'était pas encore réunie en un État unique (approximativement dans les siècles après l'an 4000). Quelques documents montrent des images pour lesquelles on hésite entre deux interprétations : dessins à intention symbolique ou véritables pictogrammes, représentations figurées équivalant à des phrases, mais non décomposables en mots. Certaines sont déjà accompagnées de mots en hiéroglyphes ; l'invention de ces caractères est donc antérieure à la période dynastique.

Débuts de la période dynastique (période thinite). — Les documents historiques, pour cette période et les suivantes, sont des listes royales reproduites imparfaitement par des auteurs grecs, retrouvées en partie dans des textes égyptiens, vérifiées par des découvertes de monuments qui peuvent être attribués avec certitude à une période, une dynastie, un prince déterminés. Un certain nombre de dynasties sont donc entrées dans le domaine de l'histoire ; les égyptologues travaillent à remplir les vides entre les périodes ainsi éclairées.

L'histoire continue commence avec Ménès (forme grécisée), premier roi de la 1re dynastie issue de la région de Thinis ; c'est le fondateur de Memphis. Son règne se situerait vers 3400, suivant la chronologie « courte » sur laquelle une partie des égyptologues tend à s'accorder ; mais d'autres le situent entre 4500 et 4000.

Pour la période de transition (2e dynastie à Thinis, 3e dynastie à Memphis), le site des pyramides de Saqqarah (3e dynastie) conserve des textes étendus en hiéroglyphes.

Ancien empire (memphite), à partir de 2900. — Les dynasties en pleine lumière de l'histoire sont la IVe (grandes pyramides de Gizeh), la Ve et la VIe.

Au cours de cette période, la langue littéraire qui s'était créée dans la période précédente s'écrit sur de nombreux monuments ; elle est maintenue sensiblement immobile pendant de longs siècles par ceux qui la manient : prêtres et scribes de profession.

L'écriture hiéroglyphique consiste en un très grand nombre

de petits dessins séparés (animaux, hommes dans différentes attitudes et parties du corps, objets divers). Dès les plus anciens textes, on trouve que certains de ces dessins n'expriment pas le nom de l'objet qu'ils représentent, mais des mots ou parties de mots homonymes à ce nom. Ce sont des hiéroglyphes phonétiques. Beaucoup équivalent à deux consonnes (les voyelles ne sont pas notées), quelques-uns à trois ; les plus fréquemment employés sont ceux qui, représentant comme idéogrammes des mots très courts, servent à noter une seule consonne. Les idéogrammes et les signes phonétiques se combinent en un système compliqué pour présenter aux yeux des lecteurs des textes continus. Les hiéroglyphes usuels sont au nombre de 600 environ. Le total des signes qu'on a dû graver pour les imprimeries bien montées, en tenant compte des diverses formes d'un même signe, oscille autour de trois mille. La direction habituelle est de haut en bas sous l'Ancien empire ; plus tard elle a été horizontale, de droite à gauche ou de gauche à droite.

L'exemple suivant (pris dans Moret, qui suit Gardiner) donne une idée de la complexité du système. C'est la légende d'une figure représentant un roi qui assomme des prisonniers.

śkr — écrit au moyen des signes suivants : *ś*, complément phonétique ; *śkr* trilitère signifiant « frapper (de la massue) » ; une massue, complément idéographique.

mntw — écrit par : *mn* ; *n* (complément phonétique du précédent) ; *t* ; *w* (désinence de pluriel masculin).

ḫa'śḫ-t — écrit par : un signe qui signifie « pays étranger », répété trois fois pour exprimer le pluriel ; *t* (suffixe de féminin).

nb — écrit par le signe *nb* « tout ».

Sens de l'ensemble : « frapper de la massue — les bédouins de — tous — pays étrangers. »

Pour reconstituer la prononciation de l'égyptien ancien, son dernier état, le copte, donne des moyens insuffisants. En particulier, pauvre en laryngales, il répond seulement par *h* et *ḫ* (dans le dialecte saïdique seulement par *h*) à quatre consonnes anciennes (qu'on transcrit *h*, *ḥ*, *ḫ*, *ẖ*). C'est en rapprochant un petit nombre de mots égyptiens de mots sémitiques de même sens qu'on arrive

à compléter les indications du copte. Grâce aux compléments ainsi obtenus, le consonantisme égyptien reçoit dans les ouvrages récents d'égyptologie un aspect très proche du sémitique, avec la même abondance de laryngales et au moins une emphatique (en tout 24 consonnes).

Moyen empire (thébain). — A partir de 2200 environ, les trois dynasties bien connues sont les XIe, XIIe et XIIIe. — Les scribes pendant cette période continuent à écrire la langue littéraire, très peu évoluée, dans les hiéroglyphes monumentaux.

C'est à cette époque qu'apparaissent pour nous les manuscrits en nombre (rares sont ceux qui datent de l'ancien empire); l'écriture en est dite, d'après le grec, hiératique, c'est-à-dire ecclésiastique, quoiqu'elle ait surtout servi à des usages profanes. C'est une manière cursive de dessiner les hiéroglyphes à l'encre, avec un roseau, sur papyrus ; les formes des hiéroglyphes ainsi simplifiés sont assez variables; le nombre des caractères de valeur différente est d'environ 600.

La langue des textes familiers est sensiblement différente de celle des inscriptions; on peut admettre que c'était l'égyptien parlé de la région de Thèbes. (Certains auteurs se refusent à employer le nom de néo-égyptien pour désigner la langue des textes hiératiques anciens.)

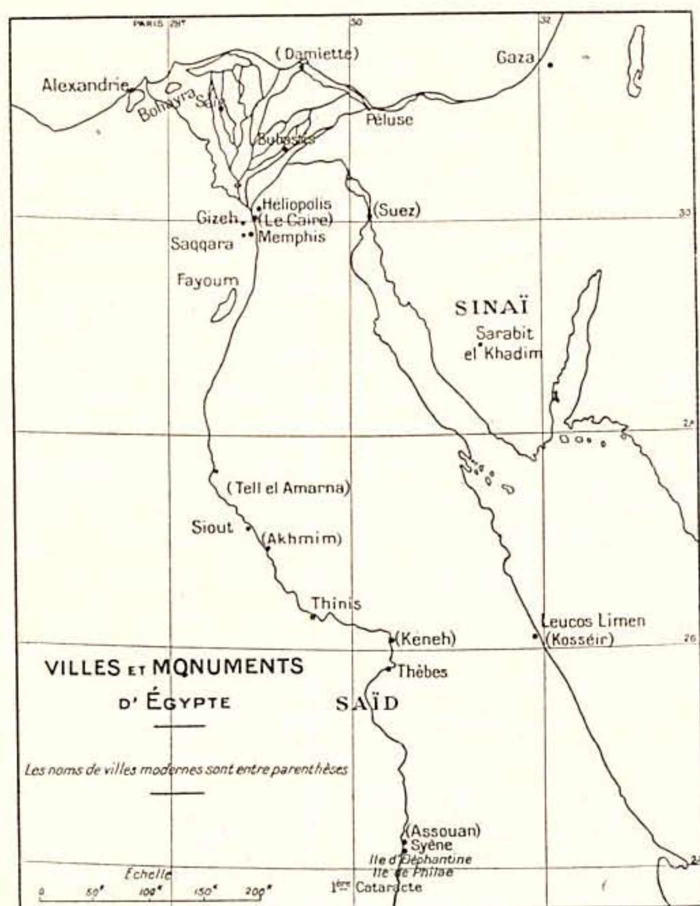
Nouvel empire (thébain). *Néo-égyptien*. — A partir de 1580 environ, les dynasties connues avec exactitude sont les XVIIIe, XIXe et XXe ; c'est l'époque des rois célèbres du nom de Ramsès.

Les scribes imitent encore la langue classique du moyen empire dans de nombreuses inscriptions.

Mais la langue vulgaire prend une place de plus en plus grande : elle est même à une certaine époque (XXe dynastie, vers 1200) écrite sur les monuments. D'abondants manuscrits sont écrits au moyen de signes hiératiques un peu évolués. Cette langue vulgaire est dite néo-égyptien.

Période saïte. — Au VIIe s. les rois de Saïs dans le Delta (les Psammétique) constituent la XXVIe dynastie, la dernière des dynasties indigènes, qui brille avant la conquête de l'Égypte par Cambyse, événement bientôt suivi lui-même de la soumission à Alexandre.

De l'époque saïte datent des inscriptions archaïsantes où l'ancien égyptien est employé de façon puriste ; c'est le fin tracé des hiéroglyphes de cette époque que reproduisent nos impressions. Des voyelles y sont quelquefois notées.



A la fin du VIe s. av. J.-C., l'écriture des hiéroglyphes sur manuscrit prenait une nouvelle forme plus cursive encore que le hiératique. On lui donne le nom de « démotique » ainsi qu'à la langue contemporaine qu'elle note, langue sensiblement une sans caractères locaux. L'écriture démotique s'essaie à noter les voyelles.

La période démotique comprend le temps des dominations perse, gréco-macédonienne (les Plolémée) et romaine, jusqu'à la victoire du christianisme.

C'est le moment où les derniers prêtres, successeurs déchus du clergé qui avait été si puissant au temps des anciens empires, perpétuaient encore à la fois la religion égyptienne et l'usage des hiéroglyphes. Ceux-ci, qui avaient été à toute époque un ornement en même temps qu'un moyen d'expression, sont employés à cette période finale de manière de plus en plus fantaisiste et ornementale, par les lettrés qui seuls désormais en connaissent encore l'usage. Déjà le démotique s'écrit en dessous des hiéroglyphes dans les inscriptions qui doivent être comprises de tous et bientôt personne n'allait plus pouvoir lire les anciens caractères ; pendant de nombreux siècles, ils sont restés lettre morte jusqu'à leur résurrection au XIX[e] s. par les efforts des savants européens (moment essentiel du déchiffrement : lecture par Champollion en 1822).

COPTE.

A partir de l'établissement du christianisme au III[e] s., les traductions des nouveaux livres saints, apportés d'abord au menu peuple par les hellénisants, s'écrivent au moyen d'une écriture elle-même grecque, dans les parlers populaires des différentes provinces.

Le nom « copte » est une forme arabe, qui répond, moins la première syllabe, au grec *Aigüptos*. En Europe on appelait le copte « égyptien » avant la redécouverte de la langue des hiéroglyphes; depuis cette découverte et avant qu'on ait distingué les différentes périodes de l'ancienne langue, le copte a été souvent nommé « néo-égyptien ».

L'alphabet copte est composé de 24 caractères grecs et de 7 lettres complémentaires prises au démotique. Les voyelles sont notées comme en grec, par une ou deux lettres.

On distingue divers dialectes en copte.

Le dialecte le plus important est le bohayrique ou boheirique,

de l'arabe *buḥayra* « province maritime » nom qui désigne le district de l'Ouest du Delta où est située Alexandrie; on le dénommait autrefois faussement « memphitique ». Ce dialecte a peut-être eu dès l'origine le caractère non d'un parler local mais du langage moyen d'une métropole à peuplement varié ; il est devenu à partir du XIe s. la seule langue littéraire et liturgique pour tous les chrétiens coptes.

En Haute-Égypte (en arabe *Sa'īd*), le dialecte important est celui de la région thébaine, qui était généralisé dans l'usage littéraire de la région avant le XIe s. ; on l'appelait autrefois thébain ; le terme préféré est maintenant saïdique, sahidique, sa'idique.

On a encore des textes dans quelques autres dialectes : le dialecte de la ville d'Akhmīm ou Akhmīn en Haute-Égypte — le dialecte de la région de Memphis — le dialecte de la région du Fayoum ou fayoumique; ce dernier dialecte avait été précédemment appelé à tort bachmourique, du nom d'un district Bachmour (*Bašmūr*, *Bušmūr*) situé dans le Delta, à l'Est de la branche de Damiette; on n'a pas conservé de textes en véritable bachmourique.

Le copte a été restreint dans sa vitalité dès le VIIe s. par l'usage de l'arabe, langue des nouveaux maîtres de l'Égypte. Parlé encore généralement par les chrétiens d'Égypte au XVe s., il n'était plus employé dès le XVIIe que par des vieillards et il était dès lors réservé à l'usage cultuel. De nos jours quelques chrétiens d'Égypte ont voulu ranimer l'usage parlé de leur langue liturgique.

Libyco-berbère.

Les Berbères, sous différents noms, couvrent ou parcourent depuis longtemps une vaste étendue, des abords de l'Égypte, qu'ils ont souvent attaquée dans l'Antiquité, jusqu'aux îles Canaries ; ils ont parfois fondé dans certaines régions des États puissants : ainsi les rois numides (surtout Massinissa) de 238 à 148 av.

J.-C., les Almoravides au XIe-XIIe s., les Almohades au XIIe-XIIIe s. Mais leur domaine consiste en partie en déserts où leur densité est très faible, en partie en massifs montagneux dont ils n'ont pas su défendre les abords contre les envahisseurs ; quant à leurs empires, ils n'ont jamais duré.

Leur langue a une histoire courte : elle n'a quasi pas donné de littérature écrite : on a seulement quelques inscriptions anciennes et, dans les temps modernes, quelques textes religieux et des textes populaires recueillis par les savants européens.

Dans l'antiquité le punique, puis le latin ont servi de langues écrites à des éléments berbères ; mais on ne sait pas quelle a été leur extension dans l'usage parlé chez les indigènes africains. Au fur à mesure de l'islamisation (VIIe siècle, puis XIe s. et suivants) l'arabe est devenu la langue écrite de presque tous les Berbères ; il s'est étendu aussi de plus en plus dans l'usage parlé et de nombreux Berbères se sont arabisés complètement, de sorte que le domaine de la langue berbère est de nos jours morcelé. Dans les dernières décades, le français a été appris par un certain nombre de Berbères (Kabyles).

Le libyco-berbère se distingue surtout par l'abondance des désinences à la fois préposées et postposées. Des exemples clairs en sont donnés par le *t* de la 2e personne dans les verbes et le *t* du féminin dans un certain nombre de noms (d'autres féminins n'ont que *t*- préfixe) ; le *t* peut s'affaiblir en *þ*, *d*, *đ* et même *h* et zéro.

Exemple en dialecte berbère des *Ait Sögruššön*, au Sud de Fez :

is	*töḫsöþ*	*a*	*töqqīmöđ*	*i*	*þöndīmþ*	*qöbbʷala*
particule interrogative	tu veux	que	tu restes	dans	(la) ville (arabe *mdīna*)	continûment (ar. *qbāla*).

« As-tu l'intention de séjourner longtemps dans cette ville-ci ? »

Dans l'état moderne de la langue l'abondance des initiales vocaliques est également remarquable. Le consonantisme, moins riche en laryngales que celui du sémitique, abonde en emphatiques. Les racines n'ont souvent, actuellement du moins, que deux consonnes (voir p. 92).

Libyque.

Le nom de Libyen, qui nous vient du grec, se trouve sous différentes formes en hébreu et en égyptien ; il n'est pas illégitime de l'appliquer à l'ensemble des indigènes de l'Afrique du Nord qui habitaient à l'Ouest de l'Égypte dans l'antiquité. Dans la région ouest de ce vaste domaine, un nom qui est attesté sous la forme grecque *Mazik(es)* se retrouve comme large dénomination ethnique moderne : *(i)mazig-ön* au Sahara et au Maroc, d'où le nom de langue *tamaziġt* ou *tamaziḫt* qui est répandu en domaine berbère, notamment dans la partie nord du Maroc.

Le libyque est représenté par quelques centaines de courtes inscriptions disséminées du Sinaï aux Canaries. Le plus grand nombre s'en trouve groupé dans les anciens domaines de Carthage ; on les a appelées aussi pour cette région numidiques ; on considère qu'aucune n'est authentiquement antérieure au IVe s. av. J.-C. ; d'une manière générale elles datent de la domination romaine. Ces inscriptions sont encore mal déchiffrées.

Il faut y joindre un certain nombre de noms propres conservés sous une forme latinisée. (Sur le *Poenulus* de Plaute, voir p. 100.)

Quelques mots qu'on a pu lire dans les inscriptions sont encore vivants en berbère ; ainsi *(a)g(ö)l(lī)d* « roi » ; d'autre part les noms propres conservés ont des formes telles qu'on est sûr d'avoir affaire à du vieux berbère ; ainsi les noms de lieu avec le double *t* du féminin : *Tubact(is)*, *tabunt(e)*.

On rencontre sur le domaine africain septentrional des graffiti que leur écriture a fait distinguer tant des inscriptions libyques proprement dites que du touareg moderne, et dont la date n'est provisoirement pas fixée ; on les appelle libyco-berbères ou sahariens.

Écritures berbères. — Le libyque est écrit dans un alphabet consonantique de 30 caractères dont les valeurs ne sont pas encore toutes connues. L'existence de bilingues punico-libyques et latino-libyques contenant abondance de noms propres permet de déterminer certaines de ces valeurs sans recourir à la comparaison avec l'alphabet moderne.

En fait, dans l'alphabet des Touaregs, qu'ils appellent *tifinag* (*tifinagh*), sur 24 caractèces (dont un signe pour les voyelles initiales), la moitié environ concorde sans difficulté avec l'alphabet libyque.

L'aspect général est resté le même : signes séparés, à formes raides, présentant surtout des formes géométriques simples telles que barre, barres parallèles, carré, cercle, etc. ; dans l'écriture moderne un point correspond souvent à une barre ancienne.

Reste à déterminer l'origine de l'écriture libyque. Il s'agit vraisemblablement d'un dérivé assez aberrant du sémitique occidental; un petit nombre de signes (six) se laisse rapprocher sans difficulté de formes phéniciennes.

Les directions de cette écriture sont très variées. Pour le libyque, la plus usuelle est de bas en haut, la colonne de droite se lisant généralement d'abord; quelques inscriptions horizontales vont de droite à gauche. Les Touaregs écrivent horizontalement, surtout de droite à gauche, mais aussi de gauche à droite ou en alternant. Mais on a signalé aussi les directions les plus variées, y compris la spirale.

Quand les Berbères emploient l'alphabet arabe pour la notation de leurs parlers, ils le complètent quelquefois en ajoutant des points à certaines lettres pour noter des consonnes qui n'existent pas en arabe.

Dialectes berbères modernes.

Le nom général de « berbère » (terme arabe) s'applique à un ensemble cohérent pour le linguiste comme pour l'ethnographe. On ne peut parler que d'une seule langue berbère, divisée en dialectes. Entre ces dialectes les différences grammaticales sont faibles. Le vocabulaire est très cohérent dans son fonds berbère, auquel s'ajoute un autre fonds commun constitué par de nombreux emprunts arabes (un petit nombre d'emprunts romans a été aussi discerné). C'est le système phonétique qui varie le plus d'un dialecte à l'autre, certains ayant plus de consonnes occlusives, d'autres plus de spirantes, ainsi *þ* au lieu de *t*. Toutes ces différences

suffisent pour que les Berbères de différentes régions ne se comprennent pas entre eux.

La distribution des dialectes berbères est exposée ici en suivant une spirale ; le groupe le moins touché d'influences extérieures est examiné d'abord, et le domaine rejeté à la fin est celui où les ilots berbères sont très espacés (voir planche 3).

La statistique est difficile encore dans la plupart des régions où le berbère est représenté ; une évaluation globale donne 6 à 7 millions de gens parlant berbère.

1. Sahara méridional. *Touareg.* — La plus grande partie du Sahara est le terrain de parcours des Berbères généralement appelés Touaregs. Ils sont restés dans leurs mœurs comme dans leur vocabulaire assez à l'abri des influences étrangères. On a évalué leur nombre à 300.000 individus.

Le mot arabe *twārög* (au singulier *targi*) reproduit un nom indigène partiel. Le nom général que se donnent les Touaregs est la forme dialectale du mot *imāzigön* (v. p. 136), soit (au pluriel) *imuhag* (sous la forme septentrionale) ou *imušag* (sous la forme méridionale), d'où le nom de langue *tamāhaq* « tamahek » ou *tamāšäq*, *tamāšägt* « tamacheq ».

Les dialectes varient légèrement suivant les régions — (du Nord au Sud) : l'Oasis de Ghat (*gat*) ; le Ahaggar, qui est le centre le plus important, celui dont la langue est seule bien étudiée (tahaggart) ; l'Aïr, l'Adrar oriental (*Adgag*), et un territoire soudanais dans la boucle du Niger.

Le domaine touareg a été entamé par l'extension du Haousa.

2. Mauritanie. *Zenaga.* — Les Zenaga qui ont conservé leur dialecte berbère sont quelques milliers d'individus dans la région de Bou Tilimmit ; ce sont les pauvres restes d'un domaine plus étendu ; le fleuve Sénégal leur doit son nom.

3. Maroc. — Le berbère est parlé dans la majeure partie des montagnes du Maroc (environ 3 millions d'individus, soit les 3/5 de la population totale).

Au Nord, les parlers rifains sont parlés dans la région méditerranéenne, sur la côte (en arabe *rīf*) à l'Est de la corne nord du Maroc ; à l'Est ils se prolongent dans les parlers de l'Algérie occidentale (voir ci-dessous).

Au centre on a souvent mis à part les parlers des Beraber (= « Berbères » en arabe), dans le Moyen Atlas au Sud de Fez. Il semble probable qu'ils ne constituent pas un groupe dialectal distinct ; les uns se rattacheraient au rifain, les autres au chleuh.

Au Sud, l'Atlas, l'Antiatlas, et la région du Sous sont occupés par le groupe des parlers « chleuh » (nom de peuple *šölḥ*, d'où le nom de langue *tašölḥīt* en berbère, *šölḥa* en arabe ; ce terme a été quelquefois appliqué à d'autres parlers berbères). La ville la plus importante du Maroc méridional, Merrakech, est en domaine arabisé ; mais elle est très fréquentée par les Berbères.

Le domaine chleuh se prolonge au Sud par la région berbère qui s'étend jusqu'à l'Oued Dra.

Au Sud-Est du Maroc, dans le Tafilelt, on rencontre des groupes de berbères appelés *qbāla* qui ont conservé leur langue au milieu de populations de langue arabe.

Une littérature religieuse en chleuh est née au moment de la réforme religieuse qui a eu pour suite la constitution de l'empire almohade au XIIe s. ; elle est malheureusement perdue. Il reste dans le même dialecte des traités religieux du XVIIIe s. et des poèmes, également religieux.

4. Algérie. — Les régions montagneuses et sahariennes ont en grande partie échappé à l'arabisation ; un essai de statistique détaillée donne (pour 1912) environ 1.300.000 individus de langue berbère (dont 725.000 ne parlant pas d'autre langue) sur un total approximatif de 4.500.000 habitants.

Il y a lieu de distinguer plusieurs régions.

Tout à l'Ouest, le prolongement des dialectes marocains du Nord se trouve en territoire algérien : Beni Snassen (*iẓnasön*), Beni Snous, Beni Bou Saïd.

Les dialectes sahariens qui se groupent avec ces parlers septentrionaux sont ceux de Figuig et de la région des Ksour sudoranais, du Gourara, du Touat et sans doute du Tidikelt.

Plus à l'Est (partie ouest du département d'Alger) un district berbère compact comprend les environs montagneux de Cherchell ; trois autres îlots plus petits et quelques parlers isolés se trouvent dans la même région.

Le correspondant méridional de ce groupe se trouve au Mzab, dont le centre principal est Ghardaia. Les Mozabites, qui sont des musulmans d'une secte spéciale, ont eu une littérature religieuse, dont il reste un traité et quelques fragments.

On peut considérer comme un prolongement oriental de la même région berbère les oasis de l'Oued Ghir (surtout Touggourt) et Ouargla.

Tous les parlers dont il vient d'être question composent l'ensemble « zénète », en prenant ce terme au sens étroit ; c'est un terme indigène qui désigne une fraction des Berbères (en berbère *ẓana*, *dẓanāt*, en arabe *ẓnātīya* « zenatia »).

A l'Est d'Alger, le berbère occupe un domaine montagneux compact : c'est ce qu'on appelle les Kabylies (grande Kabylie, et partie ouest de la petite Kabylie), sur le territoire des départements d'Alger et de Constantine ; dans cette région, par places, le berbère refoule des infiltrations arabes. Le nom kabyle est arabe : *qbāīli* « homme des tribus (*qbāīl*) ». Un des principaux dialectes est le zouaoua (*ẓwawa*, *agawawa*), parlé dans la région du Djurdjura.

Dans la partie sud du département de Constantine, la région des hauts plateaux et des monts de l'Aurès est occupée par un groupe berbère compact, auquel les Européens appliquent le nom de Chaouïa (arabe *šāwiya* « bergers »).

5. Tunisie. Tripolitaine. — La Tunisie est profondément arabisée. Le berbère n'y existe que dans le Sud : à Sened dans la région de Gafsa, dans l'île de Djerba et chez la fraction des *Mŏṭmaṭa* de *Tamŏẓraṭṭ* (Tamezred).

C'est au même ensemble que se rattache sur territoire tripolitain le groupe du Djebel Nefousa, qui touche à la côte en un point.

7. Oasis orientales. — Le domaine berbère de l'ancienne Libye est jalonné par quelques oasis : Ghadamès (*gdāmŏs*), Socna (*sokna*), Temissa, Aoudjila, et Siouah (*sīwa*), l'ancien Ammon.

Des recherches récentes permettent de croire qu'on peut diviser l'ensemble berbère en deux groupes : l'un comprendrait les par-

lers du Sud marocain (chleuh), du Sahara (touareg), de la Mauritanie (zenaga) et, dans le Nord, le kabyle ; l'autre (qui pourrait s'appeler « zénète » au sens large) comprendrait tous les autres parlers, c'est-à-dire le Nord et l'Est, moins le kabyle.

Guanche des Iles Canaries.

Le parler des Guanches, seuls habitants des îles Canaries avant l'arrivée des Européens, s'est éteint au XVIIe s. au plus tard, cédant à l'espagnol. On a comme base d'étude quelques inscriptions en caractères libyques non lues et des éléments de vocabulaire transmis par des auteurs européens depuis le XIVe siècle.

Le parler guanche se classe comme libyco-berbère par ce qu'on entrevoit de sa grammaire et par une partie au moins du vocabulaire. Il est possible qu'il ait présenté un dialecte ancien différent des dialectes de l'Afrique continentale, et d'autre part qu'il ait contenu des éléments dus à un autre parler antérieur des Iles Canaries.

Couchitique.

Les langues couchitiques occupent presque toute la corne orientale de l'Afrique jusqu'à 4° de latitude sud, en enrobant le domaine sémitique éthiopien à l'intérieur duquel elles forment encore des enclaves ; leur domaine se prolonge au Nord jusqu'en territoire égyptien entre le Nil et la mer Rouge. Dans toute la région éthiopienne, la limite occidentale de la zone montagneuse marque aussi la limite du chamitique ; au Nord seulement, le soudanais (avec le *kunama* et le *barya*) marque sur le haut plateau un éperon qui est peut-être le reste d'un domaine plus large, antérieur à l'extension du couchitique. La superficie totale approche de 2 millions de kilomètres carrés (avec sans doute environ 6 millions d'habitants), Consulter la planche 4.

Les langues qui bordent la mer Rouge, bedja, saho et afar, somali, sont les plus proches par leur morphologie du sémitique et du berbère ; celles des régions élevées, agaw et sidama, sont d'un type plus éloigné. Mais des faits de transition permettent de rattacher sans peine un groupe à l'autre : le galla, intermédiaire entre eux par sa position géographique actuelle, a surtout des rapports morphologiques avec l'agaw ; son vocabulaire le rapproche étroitement du somali et de l'afar-saho.

Les faits sont donc plus complexes que ne le supposait la division en deux groupes proposée par Reinisch (*bas couchitique*, avec bedja, afar-saho, somali, galla ; *haut couchitique*, comprenant agaw et sidama). Mais on ne peut pas encore se prononcer avec netteté sur le groupement de langues dont certaines sont encore trop mal explorées.

Malheureusement aucune des langues couchitiques n'a une littérature écrite ; elles ne sont connues que par des observations de savants européens qui datent presque toutes de la seconde moitié du XIXe siècle.

Telles qu'on les connaît, leur trait caractéristique est l'ordre des mots, qui a influencé les langues sémitiques modernes d'Abyssinie : les éléments compléments précèdent les éléments complétés ; le verbe est en fin de phrase.

Exemple (en afar) :	*alǻ*	*yō-k*	*bǻtä*	*wāh*
	chamelle	moi -à	fut perdue	je manque
anī-k	*rámili*	*yō*	*utúq*	
je suis parce que	sable	moi	jette	

« Lance-moi du sable, puisque je ne retrouve pas la chamelle que j'ai perdue. »

Pour la forme habituelle des racines (type : *bän*), voir p. 92. Le consonantisme est moins riche que celui du sémitique en éléments laryngaux ; il abonde en consonnes prépalatales.

BEDJA

Le nom auquel on donne la forme bedja d'après l'arabe écrit *bedža* est *böga* sur les anciennes inscriptions éthiopiennes ; le nom indigène de la langue est (*to*-)*beḍawye*.

Dans l'antiquité les Bedja (dont les anciens Blemmyes étaient sans doute une fraction) étaient, à l'Est, les voisins souvent hostiles des Égyptiens et des Éthiopiens de Méroé. On a pu se demander si leur langue ne se retrouverait pas dans certaines inscriptions de Méroé, y attestant leur prédominance temporaire ; rien ne l'a confirmé jusqu'à présent.

De nos jours, les *'ababde*, fraction nord des Bedia, ne sont pas seulement islamisés comme leurs congénères, mais arabisés presque entièrement pour la langue ; toutefois l'usage de leur idiome national paraît avoir encore été général chez eux au début du XIXe siècle.

Au centre on distingue les dialectes des Bichari et des Hadendoa ; au Sud celui des Halenga, dans la région de Kasala, et surtout celui des Beni 'Amer ; ceux-ci sont partagés en tribus de langue uniquement bedja, en tribus bilingues (bedja et tigré) et en tribus de langue uniquement tigré.

Le nombre des Bedja dépasse sans doute le million.

AFAR. SAHO.

Le saho et l'afar ne sont pas deux langues différentes, mais la même langue, parlée, avec certaines différences, par deux populations distinctes, encore qu'étroitement parentes. Les Saho, que les Abyssins appellent aussi Choho (*šoho*) sont des pasteurs presque tous musulmans établis dans un petit district de la région de Massaoua, sur les contreforts et sur la crête du haut-plateau abyssin. Les fractions du plateau (irob-saho) ont adopté le tigrigna, ou le parlent en même temps que le saho. Le nombre total des Saho était environ 40.000 au milieu du XIXe siècle.

Les Afar (*'afar*) musulmans s'appellent aussi *Dankal* (en arabe, pluriel *danākil*, singulier *dankali*) ; les Abyssins les nomment au Nord Taltal et au Sud Adal. Leur domaine, qui est désertique (aucune statistique sur les habitants), s'étend au Sud des Saho, entre la mer Rouge et le plateau abyssin ; la limite méridionale se trouve dans les contreforts montagneux au Sud de l'Aouache Les îles en bordure de cette région parlent aussi afar.

On n'a pas encore retrouvé le langage des *Doba*, antérieurs aux Afar et aux Saho dans la même région.

Somali.

Le nom indigène de Somali s'applique à un ensemble de tribus plus ou moins complètement islamisées qui peuplent l'extrême corne orientale de l'Afrique, à l'Est du 33° longitude Est Greenwich, de Djibouti à Kismayo.

Il n'y a pas de langue commune somali (l'arabe est répandu sur la côte chez les hommes, voir p. 119). Les parlers semblent se grouper en dialectes qui correspondent à des confédérations de tribus.

Le langage qui est généralement décrit sous le nom de somali est parlé au Nord, par les *Isāq* (habitants de la Somalie anglaise et Somali du Nord habitants à Aden).

Le dialecte le plus répandu est, au centre, celui des *Darod*, dans la province Ogaden en Abyssinie, et dans la majeure partie de la Somalie italienne ; la grande subdivision des Darod, celle des *Midžurtin* est souvent prise par erreur pour l'ensemble.

Dans la vallée du *Wēbi Šebeli* se parle le dialecte des *Hawiyya* ; on y rencontre aussi des Somali parlant galla.

Enfin dans le Sud (province Benadir) domine le dialecte de l'élément qui est nommé *Sab* par les *Hawiyya* et qui est répandu à Aden parmi les immigrants étrangers appelés Djabarti. Les Sab, ainsi que certains hors-castes du Nord paraissent des représentants plus ou moins purs d'une race antérieure aux Somali.

Chez les Hawiyya des éléments vassaux appelés « esclaves » *adōn* ont eu, à côté du somali, un langage bantou qui semble disparu.

Le somali n'est pas écrit actuellement. Les signes gravés qui couvrent certains rochers en pays somali ont été interprétés provisoirement comme des marques de passage des transhumants.

Galla.

Galla est un terme indigène partiel qui a été généralisé par les

Abyssins et transmis par eux aux Européens; les Galla eux-mêmes se nomment *oromo*, ainsi que leur langue. Cette langue a de nos jours un vaste domaine qui s'étend (du Sud au Nord) depuis la partie nord de l'Afrique orientale anglaise jusque sur le plateau abyssin dans sa partie centrale.

Il est probable qu'avant le XVe siècle les Galla s'étendaient plus au Sud ; il est certain que c'est seulement à partir de ce moment qu'ils ont occupé la partie septentrionale de leur domaine : ils ont séparé les Somali (ainsi rejetés vers l'Est) des Sidama (voir ci-dessous), recouvert une vaste région sidama (le massif montagneux d'où descendent l'Aouache, l'Omo et le Baro) jusqu'à l'Abbay au Nord, et débordé jusqu'au delà du Choa.

Les Galla sont païens dans l'ensemble. Ceux du Nord se sont en partie islamisés ; au XVIe s., ils ont été plus ou moins mêlés à l'assaut musulman contre l'Abyssinie qui leur a ouvert la voie de l'invasion. Mais ceux qui, individuellement ou en groupe, en sont venus à jouer un rôle important en Abyssinie se sont chacun à leur tour christianisés. De nos jours la religion chrétienne et la langue amharique ont repris ou reprennent possession d'enclaves galla (comme celle des *Wollo* entre le Choa et le Lasta) et des régions frontières.

Les Galla des différentes régions septentrionales paraissent se comprendre entre eux sans avoir le sentiment de différences dialectales profondes ; il semble qu'il y a cependant lieu de distinguer les parlers *maṭšṭšā* à l'Ouest, *tulama* au Nord-Est, *borana* au Sud-Est ; à cet ensemble septentrional s'oppose le dialecte méridional des Bararetta.

Le galla n'est pas écrit. Depuis le XVIe s. les Abyssins en ont noté des éléments dans leur écriture ; celle-ci a aussi quelquefois servi au XIXe s. à l'édition de textes en galla par les missionnaires européens. Une lettre en galla écrite dans des caractères par ailleurs inconnus, qui était venue en 1842 entre les mains de d'Abbadie, est restée un document isolé.

AGAW.

Agaw (*Agaga*, d'après Beke) est le nom des populations chami-

tiques qui, avant l'arrivée des Sémites, couvraient la majeure partie du haut plateau abyssin, approximativement de 10° jusqu'au delà de 14° latitude Nord. La sémitisation progressive, qui a eu pour contrepartie une certaine couchitisation des langues sémitiques, a été telle que de nos jours les parlers agaw apparaissent généralement comme des îlots linguistiques plus ou moins près de disparaître ; la plupart sinon la totalité des gens de langue agaw savent aussi l'amharique ou le tigrigna.

Les domaines séparés ont des dialectes distincts avec des noms particuliers.

1. *Bilin* (au Nord). — Les *Bilin* (*Bilen*) ou *Bogos* (*Bâqus*) sont une population musulmane actuellement peu nombreuse, dans la vallée moyenne de la rivière Anseba en Érythrée, et dans les montagnes environnantes (environ 2.000 kilomètres carrés, sans doute plus de 15.000 âmes), à la frontière linguistique du tigrigna, du tigré, du saho et du bedja. Ils sont arrivés dans cette région sans doute au xe-xie s. ap. J.-C., venant du Lasta, et occupaient autrefois une région plus vaste que de nos jours. Le bilin, reculant surtout devant le tigrigna, a gagné quelque peu sur le tigré. Il a été bien étudié au xixe siècle.

2. *Ḫamir* et *ḫamta* (au centre, vers l'Est). Les domaines de ces deux dialectes (environ 5.000 à 6.000 kilomètres carrés), à l'Est du Takkazé, sont séparés par la frontière du Tigré et du Lasta.

Le *ḫamir* (chamir, hhamara) est parlé dans le Wag, pointe entre le Takkazé et le Tsellari, et dans la région de Sokota (*Soqoṭa*), c'est-à-dire dans le Lasta septentrional. Le Lasta méridional a été le siège de l'empire chrétien d'Abyssinie pendant un siècle environ (xiie-xiiie s.), avec une dynastie agaw ; mais l'amharique semble y avoir prévalu depuis longtemps.

Le nom de *ḫamta* (khamta), terme indigène de la même racine que *ḫamīr*, s'applique aux parlers situés au Nord du Tsellari, dans un district qui paraît avoir été indépendant et non chrétien jusqu'au xvie s. environ.

Les indications de d'Abbadie sur la langue *ḫamtõña*, se rapportent à cette région de même que celles de Bruce sur les agaw tcheratz.

3. *Qwara*, etc. (au centre, vers l'Ouest). Le groupe de l'Ouest se situe, de manière discontinue, entre le bord ouest du plateau abyssin sur la latitude du lac Tana (*Sanā*) à l'Ouest et le fleuve Takkazé à l'Est.

Dans le *Qwarā* (Kouara) à l'Ouest du lac, il semble exister encore un domaine continu du parler *qwārasā*.

Jusqu'au XVIIe siècle, un État agaw indépendant, de religion juive (non orthodoxe), s'est maintenu dans le *Samēn*, vaste massif montagneux, où au moins un îlot de langue agaw a encore été signalé au milieu du XIXe siècle.

De nos jours les Abyssins de religion juive ne sont plus que des artisans dispersés en petits groupes dans les villes et villages de l'Abyssinie du Nord, de l'Ouest du lac Tana jusqu'en Érythrée ; on les appelle généralement falacha (*falāšā*) en pays de langue amharique et kayla dans le Wagara et dans le Tigré. L'usage de l'agaw sous différents noms, et peut-être avec différents dialectes (*falāšā*, *kaylōñā*, *qwarasa*) était déjà précaire chez les Falacha il y a cinquante ans ; ils parlent maintenant soit amharique, soit tigrigna. Ils emploient comme livres religieux (*Bible*, etc.) les textes guèzes, complétés quelquefois de gloses en agaw qui servent aux prêtres pour la paraphrase orale.

Les *kömant* (en amharique *qömānt* ; dits aussi kamant, gamant, etc.) sont les restes d'une population non assimilée aux Abyssins dans le Dembia (*Dämbyā*, plaine au Nord du lac Tana) et districts voisins ; baptisés, mais observant certaines pratiques juives, ils se maintiennent en élément distinct religieusement et socialement. Leur langue est dite *kömantnay*, mais rien ne prouve qu'elle soit une, ni qu'elle leur soit propre.

Une autre population des bords du lac Tana, les pêcheurs *wayṯo*, qui ont des congénères en plusieurs autres endroits de l'Abyssinie (restes probables de populations anciennes du pays) ont un langage spécial sur lequel on n'a aucune donnée.

4. *Agaw* du Sud. L'agaw est encore parlé (avec l'amharique) dans la région dite Agaoumeder (*Agāw-mödör* « pays des Agaw », en Amharique) au Sud du Qwara et du lac Tana, à l'Ouest du Godjam. Les indigènes nomment leur langage *awiya*, au moins dans la région des sources de l'Abbay.

Un dialecte très proche est situé immédiatement à l'Est de la partie sud de l'Agaoumeder, dans le Damot ; ce parler (*damōt*, *damōtōñā*), aujourd'hui d'usage très restreint, paraît avoir été bien vivant au XVIIe-XVIIIe siècle.

Sidama.

Le nom de Sidama est donné par les Galla à leurs voisins du Sud-Ouest. Ce sont dans l'ensemble les occupants de la région haute entre la rive droite de l'Abbay au Nord et les abords du lac Rodolphe au Sud, en prolongement des Agaw vers le Sud. Ce domaine a été coupé et restreint de la moitié environ par la conquête galla ; il est encore d'environ 100.000 kilomètres carrés. Les Sidama ont, avant l'arrivée des Galla, subi l'influence de l'Abyssinie sémitisée et ils avaient en partie reçu le christianisme. D'autre part ils paraissent avoir donné leur langue à un certain nombre de populations nègres qui vivaient à leur contact sur les contreforts ouest du plateau abyssin.

Les parlers sidama sont très mal connus. Ils paraissent diverger fortement entre eux. Le type chamitique y est altéré ; il semble qu'ils ont subi profondément et diversement l'influence des langues nilotiques (soudanaises). Aucun n'a donné de littérature.

L'énumération des parlers sidama est faite ici en allant du Nord au Sud.

Gonga. — Sur les deux rives de l'Abbay, à l'Ouest du Gafat, appelé aussi *šinaša* (sinacha). Le *šat* et le langage des *zēt* païens, nommés par des auteurs du XVIe-XVIIe s., étaient situés dans le même district ; ils n'ont pas été signalés par les auteurs modernes.

La langue de l'Ennarya, province anciennement chrétienne, mais submergée ensuite par les Galla, paraît avoir été au moins proche du gonga, et lui était peut-être même identique. Sa disparition a séparé le gonga du reste du sidama.

Gunza (*gwönza*). — Langue d'éléments nègres, sur les contreforts du massif abyssin, dans la boucle (et très peu sur la rive gauche) de l'Abbay ; ils sont signalés aussi sous les noms de *dizzela*, de *naga* et de *šangalla* (changalla est le nom général,

chez les Abyssins, des peuplades nègres du domaine nilotique). Le rapport du gunza avec le sidama n'est pas clairement établi.

Zendjero (djandjero, yangaro) ou yamma. — Dans l'ancien royaume de ce nom, sur la rive droite du haut Omo.

Gudella, etc. — Sur la rive gauche du haut Omo, à l'Est, surtout au Sud, et enfin au milieu même du Gouragué, quelques parlers semblent former un ensemble. On y distingue le parler du *Gudella* (*Gudela*), appelé aussi *Hadya*; celui du *Kambatta* ou *Kambat*; le district voisin *Tambaro* n'aurait pas de parler distinct.

Kafa. — Langue d'un royaume dont l'indépendance s'est maintenue jusque vers la fin du XIX[e] s., avec une aristocratie gonga. Une portion du langage kafa (kaffa, kafetcho) est restée détachée dans l'*Affilo* (*Amfilo*) comme enclave en pays galla. Le même langage est parlé aussi dans le *Garo* et le *Motša*.

Gimirra. — Le pays *Gimirra* (Ghimirra) est limitrophe du Kafa au Sud-Ouest; certains au moins des dialectes parlés par les négroïdes de cette région paraissent avoir des éléments communs avec le kafa.

Kullo-Walamo. — A l'Est du Kafa et au Sud du groupe *Gudella*, sur les deux rives de l'Omo, le sidama est parlé par diverses populations; les langages désignés par les noms de ces populations ne sont qu'une même langue, d'après plusieurs renseignements concordants. Sur la rive droite de l'Omo se trouvent les éléments dits *Wamate* (Omaté) dans le *Kullo* (Coullo) ou *Dawaro* (Dawrowa) qui est appelé *Warattä* par les Galla, dans le *Kontab* ou *Konta*, et dans le *Koyša* ou *Qutša* (Kouicha). Sur la rive gauche habitent les *Walamo* (Ouallamo), dont on a aussi appelé la langue *walaytsa*; ils occupent différents districts jusqu'au lac *Abaya* (lac Marguerite).

En prolongement du *kullo-walamo*, le *gatsamba* (faussement écrit gazamba) est parlé dans la plus grande île du lac Abaya.

Badditu. — A l'Est du lac Abaya, dans une région montagneuse, à laquelle on a donné le nom de Sidamo, sont parlés le *badditu* et le *dẓamdẓam*; ils paraissent proches du *walamo*.

Bambala. — Au Sud du *badditu*, un dialecte assez différent est parlé par les *Bambala* dont le centre est *Amarr Burdẓi* (Amara

Bourdji, Ammaro), au Sud du lac Tchamo ; leur frontière est à environ 15 kilomètres au Nord de ce point.

Marcel COHEN.

INDICATIONS BIBLIOGRAPHIQUES

Sémitique.

Carl BROCKELMANN, *Grundriss der vergleichenden Grammatik der semitischen Sprachen*. Berlin, I 1908, II 1913. Contient tout l'essentiel, avec des indications bibliographiques ; celles qui concernent des périodiques permettent de poursuivre la bibliographie pour les années suivantes.

ENCYCLOPÉDIE DE L'ISLAM, Leyde-Paris, depuis 1913 ; en particulier, article *Arabie*.

Allan H. GARDINER, The egyptian origin of the semitic alphabet. — A. E. COWLEY, The origin of the semitic alphabet. Extrait de *The Journal of egyptian archaeology*, janvier 1916.

Henri SOTTAS, Une nouvelle théorie sur l'origine égyptienne de l'alphabet sémitique. *Journal asiatique*, janvier-mars 1921.

Chamitique en général.

Carl MEINHOF, *Die Sprachen der Hamiten*. Hambourg, 1912.

Leo REINISCH, *Das persönliche Fürwort und die Verbalflexion in den chamito-semitischen Sprachen*, Vienne, 1909.

Égyptien et Copte.

Adolf ERMAN, *Aegyptische Grammatik*, 3e éd., Berlin, 1911 (bibliographie complète, comprenant les périodiques).

Alexandre MORET, *L'écriture hiéroglyphique en Egypte*, extrait de *Scientia*, février 1919 (avec bibliographie).

H. SOTTAS et E. DRIOTON, *Introduction à l'étude des hiéroglyphes*. Paris, 1922.

Georg STEINDORFF, *Koptische Grammatik*, 2e édit., Berlin, 1904,

Libyque et Berbère.

Oric BATES, *The eastern Libyans* (An Essay). Londres, 1914.

E. F. GAUTIER, *Le Sahara*, Paris, 1923.

Stéphane GSELL, *Histoire ancienne de l'Afrique du Nord*. Tome I (chapitre 5), Paris, 1913.

J.-B. CHABOT, Punica XXV, *Inscriptions punico-libyques*. Journal asiatique, mars-avril 1918, p. 259.

Publications de l'École des lettres (de la Faculté) d'Alger. Diverses monographies de dialectes berbères ; dans cette collection voir en particulier :

Tome XIV, René BASSET, *Etudes sur les dialectes berbères*, 1894.

Tome LVI, Edmond DESTAING, *Etude sur le dialecte berbère des Aït Seghrouchen* (Moyen Atlas marocain). 1920.

René BASSET, *Manuel de langue kabyle* (dialecte zouaoua). Paris, 1887.

Edmond DESTAING, *Note sur la conjugaison des verbes de forme* C^1eC^2, Mémoires de la Société de Linguistique de Paris, tome XXI. — *Note sur le pronom démonstratif en berbère*, ibid., tome XXII.

E. DOUTTÉ et E.-F. GAUTIER, *Enquête sur la dispersion de la langue berbère en Algérie*. Alger, 1913.

Archives berbères (Publication du Comité d'études berbères de Rabat), 1915 et ss. — *Bulletin de l'Institut des Hautes Etudes marocaines*, 1920. — *Hesperis*, à partir de 1921.

John ABERCROMBY, *A study of the ancient speech of the Canary islands*, dans Harvard African studies, I, Cambridge (États-Unis), 1917.

Couchitique.

Leo REINISCH, Œuvres nombreuses dans les publications de l'Académie de Vienne ou dans les collections subventionnées par cette Académie, sur le bedja, le saho, les dialectes agaw, le kafa, et en particulier *die Somali-Sprache*, 3 vol. 1900-1903.

Franz PRAETORIUS, *Ueber die hamitischen Sprachen Ostafrika's*, dans *Beiträge zur Assyriologie* II, 1894. — *Zur Grammatik der Gallasprache*. Berlin, 1893.

Enrico CERULLI, *The folk-literature of the Galla of Southern Abyssinia*, extrait de Harvard African Studies, III, Cambridge (États-Unis), 1922.

Carlo CONTI ROSSINI, *La langue des Kemant en Abyssinie*. Vienne, 1912 (avec comparaison des autres dialectes agaw). — *Studi su popolazioni dell' Etiopia*. Rome, 1914 (extrait de *Rivista degli Studi orientali*).

George MONTANDON, *Au pays Ghimirra*, Paris, Neuchâtel, s. d. (1913).

LANGUES FINNO-OUGRIENNES

ET

LANGUES SAMOYÈDES

Classification.

Les langues finno-ougriennes sont : le finnois et les langues de son groupe (carélien, olonetsien, vepse, vote, estonien, live), le lapon, le mordve, le tchérémisse, les langues permiennes (zyriène et votiak), le vogoule, l'ostiak et le hongrois. Toutes ces langues sont parentes, c'est-à-dire qu'elles continuent une même langue : le finno-ougrien commun. Ceci résulte de la comparaison des mots et des formes linguistiques d'une langue à l'autre, ainsi qu'en feront foi les exemples suivants : le nom de la « main » est dans ces différentes langues : *käsi* (fi.) *gietta* (lp.) *k'äd'* (mord.) *kit* (tchér.) *ki* (votk., zyr.) *ket* (ostk.) *kăt* (vog.) *kēz* (hong.). De même pour le nom de l'œil : *silmä* (fi.) *tšalbme* (lp.) *s'el'm'e* (mord.) *sinzä* (tchér.) *s'in(m-)* (votk. zyr.) *sem* (ostk.) *säm* (vog.) *sem* (hong.). On pourrait multiplier ces exemples.

La langue finno-ougrienne commune a été parlée par une population sans doute riveraine du cours moyen de la Volga. Elle a dû se trouver en contact avec des populations de langue indo-européenne, comme le prouvent certains emprunts très anciens. Ce furent les tribus *ougriennes* qui se détachèrent les premières de la souche commune. Elles allèrent habiter le versant occidental de l'Oural et passèrent bientôt sur le penchant asiatique. C'est là que sont restés les Ostiaks et les Vogoules alors que les Hongrois après une longue migration allèrent s'établir dans la Hongrie actuelle.

Cependant, le groupe finno-permien resté sur place s'étendait

insensiblement à l'Ouest de la boucle de la Volga. C'est alors que s'en détachèrent les tribus *permiennes* : les VOTIAKS et les ZYRIÈNES. Une dernière dislocation se produisait bientôt dans le tronçon principal resté attaché aux rives de la Volga. Tandis que les MORDVES et les TCHÉRÉMISSES restaient fidèles à leur lieu d'origine, les populations finnoises, caréliennes, vepses, votes, estoniennes et lives poussaient vers l'Occident. Elles parvenaient dans la région des lacs Onéga et Ladoga ainsi que sur la côte balte, et les Finnois occupaient vers le VIe s. de notre ère le territoire actuel de la Finlande, qu'ils partagèrent avec les Caréliens.

Quant aux Lapons, la linguistique et l'ethnographie sont bien embarrassées pour leur attribuer une origine déterminée. Ils parlent actuellement une langue manifestement finno-ougrienne mais ils ne l'ont pas toujours parlée. Peut-être faut-il voir en eux les habitants les plus anciens de la presqu'île scandinave et de la Finlande. Pour une raison inconnue, ils ont dû adopter très anciennement la langue des Finno-Ougriens, peut-être à l'époque où ceux-ci ne formaient encore qu'un seul peuple. Ce qui est sûr, c'est qu'ils se trouvaient en Finlande à l'arrivée des Finnois et des Caréliens et qu'ils ne reculèrent que lentement devant ces derniers. En 1300, on les rencontrait encore presque au centre de la Finlande.

Les langues samoyèdes et les langues finno-ougriennes sont liées entre elles. Elles remontent les unes et les autres, par delà le finno-ougrien commun et le samoyède commun, à une langue que l'on a appelée : langue ouralienne. Par là on a voulu dire que la population qui l'aurait parlée se serait trouvée cantonnée au pied de l'Oural. L'état actuel de la linguistique et de l'archéologie ne permet pas de préciser davantage.

Certains savants sont allés plus loin, ils ont prétendu rattacher aux langues ouraliennes les langues dites altaïques, sans parler du japonais. Les essais de démonstration qu'ils ont fournis sont restés jusqu'ici peu convaincants.

DIVISION DES LANGUES FINNO-OUGRIENNES

Groupe finnois. — Il comprend le finnois, le carélien, l'olonetsien, le vepse, le vote, l'estonien et le live.

Le FINNOIS (*suomi*) est parlé en Finlande, dont la plus grande partie des habitants sont de langue et de race finnoises. On le parle dans le Nord de la Suède, en Värmland, dans une enclave isolée de Norvège et parmi les immigrants finnois aux États-Unis. En tout 3.000.000 de sujets. Il comprend quatre dialectes : celui de l'Ouest, celui du Centre, celui du Sud-ouest et celui de l'Est. Il est attesté dès le XII[e] s. par des mots, des noms isolés. Le premier livre finnois a paru en 1544.

Le CARÉLIEN est parlé dans les anciens gouvernements russes d'Arkhangelsk, Olonets, Tver et Novgorod, l'OLONETSIEN dans le gouvernement d'Olonets, et en Finlande sur les bords du Ladoga, soit en tout 205.600 âmes d'après le recensement de 1897, qui ne distingue pas entre Carélien et Olonetsien.

Le VEPSE est parlé dans le gouvernement d'Olonets sur l'Oyat supérieure et sur le versant sud-ouest du lac Onéga. Le nombre des Vepses ne s'élève qu'à 25.000 environ.

Les VOTES qui sont cantonnés autour de Pétrograd sont environ 1.000. Ils sont sans doute les anciens habitants de l'Ingrie et les restes des *Tchoudes* dont parle la chronique de Nestor.

Les ESTONIENS (ou *Estes*) sont établis au nombre de 1.450.000, sur la côte baltique en face de la côte finlandaise. On les trouve encore en Livonie et, par groupes plus petits, dans les gouvernements de Pétrograd, Vitebsk et Pskov. Leur langue (*eesti*) connait deux dialectes principaux : celui de Reval et celui de Dorpat. L'*estonien* a eu une littérature assez riche, qui a commencé à se développer au XVII[e] siècle.

Les LIVES (*randalist* ou *kala'mi'ed*) habitent 12 villages de la côte nord de la Courlande. Leur nombre était en 1888 de 3.000 environ, mais ils ont dû diminuer beaucoup depuis. Ils sont fortement lettisés.

Les LAPONS (*sābme*, *sābmelaš* ou *sāmiladš*) sont répartis entre la Norvège, la Suède, la Finlande et la Russie. Leur nombre total est de 28.500 environ. On distingue 5 dialectes lapons : celui parlé en Norvège (lp. N.). celui qui se parle en Suède (lp. S.) celui qui se parle en Suède également dans la région de Luleå (lp. L.), celui parlé en Finlande à Inari (lp. I.), celui de Kola, en Russie, sur la côte de la mer Blanche (lp. K).

Le MORDVE est la langue des populations établies sur le cours moyen de la Volga (Samara, Simbirsk, Penza, Saratov, Tambov et Nijni Novgorod). Ces populations comptaient 1.000.000 d'âmes environ en 1897. Il y a deux dialectes : le mordve *erzä* et le mordve *mokša*.

Les TCHÉRÉMISSES (*mari*) sont dispersés principalement dans les gouvernements de Viatka, de Kazan et d'Oufa, moins dans ceux de Perm, de Kostroma et de Nijni Novgorod. Ils étaient 375.000 en 1897, dont 28 °/₀ païens. Ils parlent trois dialectes selon qu'ils sont de la montagne, de la prairie ou de l'Est.

Les ZYRIÈNES (Syriènes) (dans leur langue : *komi*) s'étendent sur tout le territoire arrosé par la Pétchora, l'Ijma, la Mézéne, la Vachka, la Vytchegda, la Sysola, la Louza et la Kama (gouvernements de Vologda, Arkhangelsk, Viatka et Perm). Ils étaient 258.000 en 1897 et forment actuellement une république autonome. Leurs dialectes sont au nombre de 10. Le zyriène est attesté dès la deuxième moitié du XIVe siècle.

Les VOTIAKS (*ud-murt*), très proches parents des Zyriènes forment avec eux le groupe permien. Ils habitent la région comprise entre la Viatka et le Kama (gouvernements de Viatka et d'Oufa.). On en compte 420.000, toujours d'après les recensement de 1897. Les Bessermaniens forment un petit groupe à part. Ce sont des Turcotatares qui ont adopté pour langue le votiak. Outre le dialecte bessermanien, le votiak se divise en 7 dialectes.

Le VOGOULE est la langue d'une population en voie d'extinction.

Les peuplades vogoules (*men'dši-* ou *mån'ši*) dont l'ensemble n'atteint pas 5.000 âmes sont réparties sur les deux versants de l'Oural, entre les gouvernements de Perm et de Tobolsk. Il y a deux zones dialectales : celle du Nord (dialectes de la Lozva supérieure, de la Lozva et de l'Ob) et celle du Sud (dialecte de la Lozva moyenne et inférieure, de la Pelymka, de la Konda et de la Tavda).

Le domaine OSTIAK s'étend à l'Est du territoire vogoule sur l'Ob et ses affluents, dans le gouvernement de Tobolsk. Les Ostiaks (*handa-hui*) qui étaient 19.000 en 1897 pratiquent tout comme les Vogoules le chamanisme. Leur langue connaît 7 dialectes et est très parente du vogoule avec laquelle elle forme le groupe des langues dites ob-ougriennes, c'est-à-dire langues ougriennes de l'Ob.

Le HONGROIS (*måd'år*) est celle de toutes les langues finno-ougriennes qui est parlée par le plus grand nombre de sujets. Il est la langue de 9 1/2 millions d'âmes. Hors de la Hongrie proprement dite, le hongrois est parlé dans les colonies hongroises de Roumanie, Bukovine, Transylvanie, etc., de Yougoslavie et d'Autriche (Burgenland.). Les linguistes hongrois partagent le domaine de leur langue en 8 dialectes principaux : 1) le dialecte de l'Ouest, 2) celui parlé au delà du Danube, 3) le dialecte de l'Alföld, 4) le dialecte de la région Tisza-Danube, 5) celui du Nord-Ouest, 6) celui du Nord-Est, 7) le dialecte au delà du Königsteig et 8) celui des Szekler auquel se rattachent les parlers des colons de la Bukovine et de la Roumanie. Tous ces dialectes diffèrent du reste assez peu les uns des autres.

Le hongrois se rattache aux langues ob-ougriennes avec lesquelles il forme le groupe *ougrien*. Ces trois langues se distinguent des autres idiomes finno-ougriens par un ensemble de particularités communes.

Le hongrois est de toutes les langues finno-ougriennes, celle qui est le plus anciennement attestée. Outre quelques mots isolés (au nombre d'une cinquantaine) dans un document de l'abbaye de Tihany (1055), l'on possède une oraison funèbre du XIIIe siècle et quelques fragments datant du XIVe.

LANGUES SAMOYÈDES

Au nombre de 18.000 seulement, les Samoyèdes vivent dispersés sur une étendue immense, depuis les contreforts européens de l'Oural, jusqu'au delà de l'Iénisseï, des monts Sayan jusqu'à la Mer de glace. Les plus nombreux, les Youraks, errent en nomades sur la toundra entre Arkhangelsk et l'embouchure du fleuve Iénisseï. Plus loin vers l'Est errent, autres nomades, les Samoyèdes de l'Iénisseï et de l'Avam, dont les tribus sont en voie de disparition, et entre l'Iénisséï et la baie de Khatanga les Samoyèdes tavguis.

Entre l'Ob et l'Iénisséï et dans les forêts qui séparent les uns des autres les gouvernements de Tomsk, de Tobolsk et de Iénisséi, on rencontre les Samoyèdes ostiaks, qui sont environ 3.000 et ne se répartissent pas en moins de 20 dialectes différents. Bien plus au Sud, sur le versant nord des monts Sayan subsiste un petit reste des tribus Kamassiques, autrefois assez nombreuses. L'ensemble du groupe samoyède comporte donc cinq langues nettement différenciées, riches de plus de 40 dialectes. Ce sont : le samoyède ostiak, le samoyède yourak, le samoyède kamassique, le samoyède tavgui et le samoyède de l'Iénisséi.

Le type linguistique initial.

Structure phonétique.

La linguistique comparée des langues finno-ougriennes et des langues samoyèdes est parvenue à restituer dans ses traits essentiels la structure phonétique de l'idiome initial.

Contrairement à ce qui avait été longtemps admis, les dernières recherches tendent à prouver que l'ouralien commun ne connaissait pas un accent fixe. La place de l'accent qui est sur la première syllabe du mot en finnois, en hongrois, en lapon, en vogoule n'est pas à beaucoup près un fait aussi caractéristique que certains savants

avaient pensé. Ni le dialecte tavda du vogoule ni le votiak ne placent l'accent en tête du mot. Bien plus, en zyriène, sa place varie selon les dialectes. En tchérémisse c'est généralement la dernière ou l'avant-dernière syllabe qui est accentuée. A l'origine l'accent tombait alternativement sur la première et la deuxième syllabe. C'est aussi ce que semble montrer le phénomène de l'alternance consonantique et vocalique, ainsi que la chute de la voyelle à l'initiale de certains mots ; le hongrois a en effet : *nåpå* « belle-mère », en regard de fi. *anoppi*, vog. ostk. *ånep*.

M. E. N. Setälä a démontré l'existence en finno-ougrien commun (et en samoyède commun) d'un système d'*alternances* consonantiques (*Finnisch-ugrische Forschungen* XII, *Anzeiger*, p. 1 et suiv.). Les consonnes intérieures d'un même mot pouvaient apparaître sous différents aspects, selon qu'elles présentaient le degré fort ou le degré faible. La place de l'accent aurait entraîné l'un ou l'autre degré selon les cas. Ainsi, en indiquant par ` la quantité moyenne de la consonne, on aurait eu les alternances représentées, par exemple, dans les schémas suivants :

DEGRÉ FORT	DEGRÉ FAIBLE	DEGRÉ FORT	DEGRÉ FAIBLE	DEGRÉ FORT	DEGRÉ FAIBLE
k`k	*k*	*k*	*g*	*ňk*	*ňg*
t`t	*t*	*t*	*đ*	*nt*	*nd*
p`p	*p*	*p*	*ƀ*	*mp*	*mb*

Les langues du groupe finnois et le lapon qui ont seules conservé des restes importants de ce système peuvent encore en donner une idée.

Le finnois oppose :
nominatif *kukka* « fleur » à génitif *kukan* « de la fleur » ;
la 3e personne : *ottā* « il prend » à la 1re *otan* « je prends » ;
nom. *seppä* (forgeron) à gén. *sepän* ;
la forme *lukē* « il lit » à *luen* (< **lugen*) « je lis » ;

nom. *kota* « cabane » à gén. *kodan* (*< *kođan*);
nom. *tapa* « habitude » à gén. *tavan* (< **taban*);
nom. *laňka* (fil) à gén. *laňgan* ;
antä « il donne » à *annan* (< *andan*) « je donne » ;
ampū « il tire » à *ammun* (< **ambun*) « je tire ».

Le degré fort apparaissait, semble-t-il, après syllabe accentuée, le degré faible après syllabe inaccentuée. L'alternance avait lieu tantôt entre consonne longue et consonne moins longue, tantôt entre consonne forte et consonne faible. Une alternance *kk* ~ *k* est du premier type, une alternance *k* ~ *g* du second. L'alternance consonantique était donc à la fois *quantitative* et *qualitative*. Ce qui différenciait le degré fort du degré faible de consonantisme, c'était une articulation plus énergique et une fermeture plus marquée lors de la production du son. De plus le degré fort se distinguait souvent par l'absence de vibrations glottales.

Parallèlement, M. E. N. Setälä a indiqué que le finno-ougrien commun connaissait l'alternance vocalique. La syllabe accentuée comportait le degré fort de vocalisme, la syllabe inaccentuée le degré faible. Cette alternance était qualitative et consistait en une différenciation du timbre de la voyelle. Certaines langues finno-ougriennes en ont encore conservé quelques traces. Le finnois oppose : *pala-* « brûler » (intransitif) à *poltta-* « brûler » (transitif). De même en ostiak on a : *p'àrt-*, « commander » en face de *p'irtəm* « j'ai commandé », et en hongrois : *hâzâ*, « vers la maison » en face de *honn* « à la maison ».

On a souvent cité comme trait caractéristique des langues finno-ougriennes l'*harmonie* ou *assimilation vocalique*. Si un mot a pour voyelle radicale l'une des voyelles prépalatales *e*, *i*, *ä ö*, *ü*, les voyelles des syllabes qui suivent ne peuvent avoir que les mêmes timbres :

hongrois : *äpä* « fiel », *küsöb* « seuil », *d'öker* « racine »
finnois : *vesi* « eau », *käsi* « main », etc.

Si au contraire, la voyelle radicale est l'une des voyelles postpalatales *a*, *ā*, *o*, *ō*, *u*, *ū*, les voyelles suivantes ne peuvent avoir qu'un timbre postpalatal :

hongrois : *fålu* « village », *hosszu* « long », *hālō* « filet » ;
finnois : *paha* « mauvais », *mato* « ver », *muna* « œuf ».

Mais cette règle n'est pas rigoureuse. Ainsi le hongrois est plein de mots comme : *hibå* « faute », *vēkon'* « mine », *kīd'ō* « serpent » et le finnois de mots comme : *hiekka* « sable », *heikko* « faible », *talvi* « hiver », *onni* « bonheur », *lukē* « il lit », etc., qui comportent à la fois des voyelles prépalatales et postpalatales.

Dans la mesure même où elle existe, cette *harmonie vocalique* ne caractérise que quelques-unes des langues finno-ougriennes. Elle est bien développée en hongrois, un peu moins en finnois et dans le tchérémisse des montagnes, le vogoule de la Tavda et de la Basse Lozva ; mais on n'en peut découvrir la moindre trace ni en lapon, ni dans les langues permiennes, ni en ostiak.

L'étude des faits d'assimilation vocalique dans les suffixes permet de tirer cette conclusion que l'*harmonie vocalique* ne s'est développée que tardivement dans les langues du groupe ouralien et ne saurait passer pour une de leurs caractéristiques.

Structure morphologique.

La langue finno-ougrienne commune, et la langue samoyède commune ne semblent pas avoir connu de distinctions nettes entre les notions de nom d'une part et de verbe d'autre part. En ce qui concerne le nom, les notions d'adjectif et de substantif paraissent avoir également été assez peu différenciées.

Le substantif. — Contrairement à ce qui se passe en indo-européen et en sémitique, il n'y a pas de genre dans les langues ouraliennes. En revanche le vogoule, l'ostiak, le lapon et le samoyède ont maintenu un duel à côté du singulier et du pluriel.

Il y a plusieurs thèmes pour chaque nom. Le thème nominal du singulier sert de nominatif singulier. A cela s'ajoutent divers thèmes de pluriel, qui s'opposent deux à deux dans chaque langue.

Ainsi le finnois oppose le nominatif-accusatif pluriel : *linnut* « oiseaux » aux cas obliques du pluriel qui sont formés sur un thème *lintui-*. Dans ce premier cas le suffixe est *-t*. Exemples :

lapse « enfant » *lapset* « enfants » ;

kala « poisson » *kalat* « poissons » *kalain* (* < *kala-đ-en*) gén. plur. : « des poissons ».

De même en vogoule : *āmpət* « chiens », en regard de *āmp* « chien »

D'autres langues comme le hongrois et le lapon ont à la place de ce suffixe en *-t* un suffixe en *-k*.

hongrois : *hâiō* « navire », *hâiōk* « navires » ; *hāz* « maison », *hāzâk* « maisons » ; *seg* « ongle », *segäk* « ongles ».

Dans le second cas le suffixe est *-y*. C'est ce que présentent les formes de pluriel finnoises :

lintui- « oiseaux » en regard de *lintu* « oiseau » ;

solmui- « nœuds » en regard de *solmu* « nœud ».

De même en hongrois : *hâiō* « navire », *hâiōi* « navires » *d'ürü*, *d'ürü-i* « anneau, anneaux » et en lp. S. : *namma* « nom » *namāi-* « noms »

Le vogoule donne (au lieu de *y*) des exemples de *-n* s'opposant au *t* :

sun « traineau » *sunan* « traîneaux »

näm « nom » *nämän* « noms »

Pour le duel le suffixe thématique est *-*g* :

vogoule *āmp* « chien » *ampig* « deux chiens » ;

ostiak *sèm* « œil » *sèmgən* « les deux yeux » ;

samoyède ostiak *kule* « corbeau », *kuleag* « les deux corbeaux » ; *hay* « œil », *hayōg* « les deux yeux ».

Toutes ces formes de duel sont des nominatifs.

Le nombre des cas de la déclinaison est très variable selon les langues. Le finnois en compte 16, l'estonien son proche parent seulement 11, le hongrois 22, le lapon 8. Les suffixes casuels ne sont pas les mêmes d'une langue à l'autre.

Tandis que le finnois marque le cas *inessif* (indiquant ce qui est placé à l'intérieur de l'objet) par le suffixe *-ssa* (< **-sna*) : *mā-ssa* « dans le pays » (*mā* « pays ») et que le mordve dit : *on -sne* « en rêve » (*on* « rêve »), le hongrois se sert d'un autre

suffixe : *-bän* ~ *-bån* :

kēz-bän : « dans la main » (*kēz* « main »)

hāz-bån : « dans la maison » (*hāz* « maison »)

On pourrait multiplier les exemples. Ceci fait comprendre que les suffixes casuels des langues ouraliennes sont pour la plupart d'anciens adverbes autonomes, employés en rapport d'annexion avec le substantif qu'ils déterminent. Ils ont fini par se souder et faire corps avec le mot. Il est né de ce fait toute une déclinaison. Mais tout révèle que cette agrégation des noms et de leurs postpositions est récente. Ainsi, le finnois construit des suffixes casuels de la façon suivante :

substantif + suffixe casuel + suffixe possessif.

tandis que le hongrois construit :

substantif + suffixe possessif + suffixe casuel.

hongr. : *hātå-m* « mon dos » (*-m* « mon ») *håtå-mo-n* « sur mon dos » (*-mo-* « mon » et *-n* « sur »)

finnois : *kirya-ssa* « dans le livre » (*-ssa* « dans »), *kirya-ssa-ni* « dans mon livre », (*-ni* « mon »).

En réalité la langue commune ne semble avoir connu que quatre suffixes casuels vraiment attestés pour l'ensemble de la famille : un locatif en **-n*, un ablatif en **-t*, un datif en **-k*, un accusatif en **-m*, qu'on retrouve également en samoyède.

Mais même ces suffixes, à l'exception peut-être de celui de l'accusatif, ne rappellent que de très loin les désinences casuelles de l'indo-européen. Ils sont plus mobiles, se joignent à n'importe quel thème, servent également au singulier, au pluriel et au duel. Le finno-ougrien ne connaît pas de catégories flexionnelles, entre lesquelles les différents substantifs ou adjectifs se répartiraient comme en indo-européen. Une opposition du type *pater*, gén. *patris*, et *dominus*, gén. *domini*, n'existe pas.

Comment le substantif finno-ougrien ou samoyède s'emploie-t-il ? Un procédé très primitif est celui de la composition par simple juxtaposition. Un thème nominal est collé comme premier terme à un autre thème nominal lequel est seul décliné. Ce deuxième terme est souvent un ancien thème verbal qui s'est vu ajouter un suffixe de dérivation nominale. C'est le type de

hongr. : *fâ-vāgō* « bûcheron » (*fâ-* « bois », *vāgō* « coupant, qui coupe ») ; de même on a *būn-bānō* « repentant » (*būn-* « péché », *bānō* « regrettant, qui regrette »). On notera que les thèmes *fâ-* et *būn-* sont employés *nus*, c'est-à-dire sans le suffixe de l'accusatif que l'on attendrait.

Aux prépositions si fréquentes dans nos langues, répondent des postpositions : dans le finnois : *ūnin takā* « de derrière le poêle », c'est *ūnin* qui a le sens de « poêle », et c'est *takā* qui correspond à notre préposition. En finnois la plupart des postpositions se construisent avec le génitif, comme dans l'exemple ci-dessus : de même dans : *īsän vieressä* : « à côté du père » (*īsän* gén. de *isä* « père » et *vieressä* « à côté »).

En hongrois par contre, et ceci répond à un usage ancien, c'est le thème *nu* qui, la plupart du temps, précède la postposition : *töld'äk ålått* « sous le chêne » (*töld'äk* « chêne », *ålått*, « sous »).

De même en vogoule : *kȣol kiȣert* : « dans la maison » (*kȣol* « maison », *kiȣert* « dans »).

Cependant, dans certaines langues, les prépositions existent. C'est surtout le cas du finnois.

On notera en passant que la plupart des postpositions des langues actuelles ne sont pas autre chose que des substantifs à un cas adverbial : le finnois *vieressä* (à côté) n'est que le cas *inessif* (indiquant ce qui est placé à l'intérieur de l'objet), du mot : *vieri* : « bord », « côté », qui existe dans la langue indépendamment de son emploi comme postposition.

Ceci amène à parler de l'usage des cas. La plupart des suffixes casuels de la déclinaison des langues finno-ougriennes et samoyèdes répondent à des périphrases adverbiales, à des prépositions ou des locutions prépositives dans les langues indo-européennes. Ainsi : *avec les yeux* se dit en finnois *silmi-llä* ;
le suffixe *-llä* (*-lla* dans les mots qui ont une voyelle radicale de timbre *a*, *o*, *u*) est appelé suffixe de l'adessif. Le mot *tie* « route » employé à l'adessif « *tiellä* signifiera » sur la route. L'adessif correspond donc aux locutions prépositionnelles françaises introduites

par *sur* et *avec*. Le suffixe casuel, disions-nous plus haut, est le même au singulier et au pluriel :

hongr. : *hàlō-bå* « dans le filet »
hālōk-bå « dans les filets ».

Il suffit de joindre le suffixe casuel au thème de pluriel : *hālōk* pour obtenir l'*inessif pluriel*.

D'ailleurs le thème nu, c'est-à-dire dépourvu de tout suffixe, n'est pas une abstraction comme en indo-européen ; il servait à l'origine de nominatif singulier. Ces thèmes étaient généralement dissyllabiques ou polysyllabiques. Ils se terminaient par une syllabe brève. C'est ce que reflètent encore aujourd'hui les nominatifs singuliers suivants :

finnois *silmä* « œil », mordve *s'eľm'ä s'eľm'e*, lp. S. *tšal*^{a}*mē*, lp. N. *tšål*^{a}*bmi*, lapon K. *tšälme* ou encore : finnois *lintu* « oiseau » ; *nimiä* « belle-fille » *veri* « sang » (< **vere*.)

La possession est indiquée par des suffixes possessifs qui se juxtaposent selon les langues soit après le thème nominal et avant le suffixe casuel soit inversement après le suffixe casuel qui est attaché dans ce cas directement au thème nominal (voir p. 163).

hongrois *kår* « bras »
kåro-m « mon bras »
kåro-d « ton bras »
kår-yå « son bras »

finnois *koira* « chien »
koira-mi « mon chien »
koira-mme' « notre chien »
koira-nne' « votre chien »
koira-nsa' « leur chien ».

Ces suffixes possessifs n'ont pas dû exister sous cette forme dans la langue initiale. Ils devaient être au contraire assez indépendants et seulement rattachés au mot par un lien très lâche. De plus, ils avaient une marque du pluriel, qui a disparu au cours des temps et dont le suffixe comportait un *-n-*. Le vogoule semble avoir gardé les vestiges les plus clairs de cet ancien état

de chose, encore qu'ils soient fort dénaturés. Si l'objet possédé est au singulier et son possesseur au singulier et de la première personne (*mon*), on a, selon les dialectes, les formes suivantes du suffixe possessif : *-em*, *-m*, *ım*, *-um* et si l'objet possédé est au pluriel (*mes*) : *-änem*, *-ām*, *-ām*, *-anum*.

Une autre particularité, c'est que la postposition ne se construit jamais avec un pronom personnel, elle prend à sa place le suffixe possessif correspondant : *à côté de moi* se dit : *vieressä-ni*, en finnois. Or *vieressä* n'est, on l'a vu, que le cas *inessif* du substantif *vieri* « côté, bord ». Ceci revient donc à dire : *à mon côté* au lieu de : *à côté de moi*.

L'adjectif. — Adjectif et nom s'entremêlent volontiers. Extérieurement, rien ne permet de distinguer l'un de l'autre. Certains suffixes de dérivation sont adjectifs, d'autres fournissent des substantifs. C'est là un point de repère. On reconnaît en outre l'adjectif à ce qu'il est invariable et précède le nom qu'il qualifie, tout comme en anglais moderne. Son emploi avec le nom forme comme une sorte de composition à deux termes et c'est ce qui explique que le génitif finnois puisse provenir d'un ancien adjectif dérivé, indiquant la possession : de *valtio* « État », on formait l'adjectif, *valtion*, « qui appartient à l'État », d'où l'expression finnoise : *tämä metsä on valtion* « cette forêt est de l'État », c'est-à-dire « appartient à l'État ».

Seuls le groupe finnois, le lapon et le hongrois sont parvenus à développer un comparatif de l'adjectif.

fi. : *vanha* : « vieux »
vanhempa- « plus vieux »
est. : *paha* « mauvais »
pahemB « pire »
vana « vieux »
vanhemB « plus vieux »
lp. K. : *kieppes* « léger »
kiepsamp « plus léger »
hongr. : *nåd'* « grand »
nåd'obb « plus grand »

rövid « court »
rövidäbb « plus court »

Encore ce comparatif peut-il également s'appliquer à un substantif :

finnois *ranta* « bord, rive », *rannempana* « plus près du bord »
pohya « Nord » *pohyemmaksi* « plus loin vers le Nord » (le suffixe *-ksi* correspond à notre préposition *vers*).

Aussi Heinrich Winkler et M. Yrjö Wickman ont-ils voulu voir dans ces formes de comparatif des vestiges d'anciennes dérivations intensives. C'est ce qui explique des emplois comme : finnois *lapsi* « enfant », *lapsempa-na* « étant plus petit enfant ».

Dans la plupart des langues finno-ougriennes, le comparatif est exprimé par des moyens syntaxiques. C'est le cas ablatif qui indique l'idée de comparaison ; ainsi le zyriène dit : *meis' vilin* « plus haut que moi » (haut par rapport à moi, à partir de moi).

Quant au superlatif, seules en possèdent un les langues finnoises et le lapon. Dans les autres, ce sont des adverbes qui renforcent de manière appropriée la valeur de l'adjectif.

Le verbe. — Sans le témoignage des langues samoyèdes, on aurait peine à se faire une idée du verbe de la langue initiale. Heureusement leur témoignage permet de restituer le système verbal primitif.

Une distinction domine toute l'histoire de la conjugaison ouralienne : celle entre la conjugaison subjective et la conjugaison objective.

En samoyède, comme en hongrois, en vogoule, en ostiak et même en mordve, les formes verbales ne sont pas les mêmes selon que le verbe est transitif ou intransitif. Lorsqu'un Samoyède veut dire : « je suis homme », il s'exprime de la façon suivante :

homme-moi

et de même :

homme-toi
homme-lui

Ce que nous exprimons par l'emploi du verbe « être », il le rend par la simple juxtaposition du pronom personnel au substantif. De même veut-il indiquer qu'il est grand, il dit :

grand-moi

S'il s'agit d'une notion que nous exprimons par l'usage d'un verbe : *je vis, je marche, je vois, je coupe*, il dit encore : *vivre-moi, marcher-moi, voir-moi, couper-moi*. Il se sert à cet effet d'un *thème verbal nu*, auquel il suffixe le pronom personnel. Ceci pour exprimer l'action *intransitive*. Cette façon de conjuguer est ce qu'on a appelé la *conjugaison subjective*. C'est en somme, sous forme d'un schéma :

thème verbal + pronom personnel *sujet*.

Une action peut avoir un objet, deux objets ou plusieurs objets. Je puis dire : je vois un cheval, je vois deux chevaux, je vois plusieurs chevaux. Dans ce cas, le Samoyède reflétant l'état ancien de la langue, ne dit pas : *voir-moi*, il dit : *voir-mon*, c'est-à-dire : *mon fait de voir*. En d'autres termes il attache au *thème verbal* non plus le pronom personnel sujet mais le *suffixe possessif*.

C'est ainsi que l'on a en samoyède tavgui.

kula-*ma* « *mon* corbeau »
kula-*ra* « *ton* corbeau »
kula-*du* « *son* corbeau »

et de façon semblable le verbe *transitif* :

matuyua-*ma*
matuyua-*ra*
matuyua-*du*

qui signifie : j'ai coupé, tu as coupé, il a coupé, (mon fait d'avoir coupé, ton fait d'avoir coupé, son fait d'avoir coupé), étant bien entendu qu'il s'agit d'une action transitive (avoir coupé *quelque chose*).

De même en samoyède yourak, à côté de :

lamba-*u* « *mon* ski »
lamba-*r* « *ton* ski »
lamba-*da* « *son* ski »

on a :

madā-*u*
madā-*r*
madā-*da* :

« je coupe, tu coupes, il coupe (mon fait de couper, ton fait de couper, son fait de couper) ».

De même encore en samoyède du Iénisséï à côté de :

l'ibe-*bo* « *mon* aigle »
l'ibe-*lo* « *ton* aigle »
l'ibe-*ra* « *son* aigle »

on a :

mota-*bo* « je coupe »
mota-*lo* « tu coupes »
mota-*ra* « il coupe »

On pourrait multiplier ces exemples empruntés à un article de Heinrich Winkler (*Finnisch-ugrische Forschungen* XIII p. 128 et suiv.).

Les choses ne sont pas restées aussi simples partout. Le hongrois dit :

kåpom « je prends »
kåpod « tu prends »

mais dans

kåpyå « il prend »
kåpyuk « nous prenons »
kåpyātok « vous prenez »
kåpyōk « ils prennent »,

un élément nouveau s'introduit, c'est l'élément : -*y*-. Ce -*y*- est une sorte d'indice pronominal, intercalé entre le thème verbal et le suffixe possessif et qui tient lieu de l'objet qui subit l'action considérée.

Quand le hongrois dit :

å kålåpot kåpyå « il prend le chapeau »

en réalité il s'exprime à peu près comme le français qui dirait : le chapeau, il *le* prend. Ce n'est pas par hasard que -*y*- a été introduit ici dans la forme verbale, car cet infixe n'est autre chose qu'un vestige du pronom relatif.

De même, lorsqu'en mordve mokša on dit : *soda-t-ä* « je te connais », ce -*t*- est le pronom personnel de la seconde personne du singulier.

Dans la plupart des langues finno-ougriennes, c'est l'autre conjugaison, la conjugaison subjective, qui s'est généralisée.

Le mordve dit : *at'a-n* « vieillard-*moi* »

at'a-t « vieillard-*toi* »

pour exprimer : je suis un vieillard, tu es un vieillard. Il met le nom au pluriel et obtient :

at'a-ta-ma « vieillard-*nous* »
at'a-ta-da « vieillards-*vous* »

c'est-à-dire : nous sommes des vieillards, vous êtes des vieillards.

Le finnois dit *kiryoita-n* « j'écris » qui est un verbe transitif en même temps qu'intransitif, bien que la désinence *-n* (< **m*) ait été à l'origine la désinence de la conjugaison subjective, c'est-à-dire uniquement intransitive.

De là au hongrois : *āllo-k* « je me tiens debout » (se tenant debout-je), *nëze-k* « je regarde » (regardant-je), il n'y avait qu'un pas. Le *-k* est le reste du pronom personnel : « je, moi » renforcé par la particule emphatique *-k-*.

La conjugaison subjective s'est donc développée en hongrois, en vogoule, en ostiak, en mordve parallèlement à la conjugaison objective. Dans les autres langues, les deux conjugaisons se sont confondues. Ce sont les formes de la conjugaison subjective qui se sont généralisées.

Même dans les langues qui opposent conjugaison objective et conjugaison subjective, les deux types de conjugaison se sont mêlés. Des formes possessives ont pénétré en grand nombre dans la conjugaison subjective, jusqu'à la rendre méconnaissable.

Le verbe connaît deux sortes de thèmes : les thèmes modaux et les thèmes temporels. Les thèmes temporels sont improprement désignés ainsi. A l'origine et presque jusqu'à nos jours, ils n'ont pas indiqué des notions de temps mais seulement d'*aspect*. Un thème exprime le procès qui est en cours de développement, un autre le procès qui parvient à son point d'achèvement. C'est ce qui explique que, sauf le hongrois, dans certains cas, aucune de ces langues ne soit parvenue à développer un futur.

Les thèmes imperfectifs et perfectifs, qui ont donné dans la suite les présents et les prétérits, étaient de véritables substantifs, traités comme tels. Ainsi à la 3ᵉ personne où le thème s'employait sans aucun suffixe, il prend la forme du pluriel régulier des substantifs pour servir de 3ᵉ personne du pluriel :

finnois : *sā* « il reçoit », *sāva-t* « ils reçoivent » ; *sāi* « il reçut », *sāiva-t* « ils reçurent »

ostiak : *yogotse-t*, vogoule : *yoxtse-t* « ils vinrent »

Même le participe passé finnois est employé dans certains cas comme substantif et se voit annexer le suffixe possessif. Ainsi le participe passé *tul-tu*, « venu », peut se mettre au cas ablatif et recevoir le suffixe possessif :

tul-tu-a-ni, « après ma venue ».

Les thèmes de présent sont des thèmes dérivés. Ils se composent d'un thème verbal auquel s'adjoint un suffixe de dérivation, lequel est *nominal*.

La langue commune semble n'avoir eu qu'un seul suffixe de présent, le suffixe en *-k-* et *-g-*.

vogoule : *min-*, *mən-* « aller » :
minèg-m
məngə-m « je vais » (allant-moi)

estonien : *sure-* « mourir », *anu-* « implorer »
surek-se « il meurt » (mourant-lui)
anuk-se « il implore » (implorant-lui)

Le suffixe de prétérit général aux langues finno-ougriennes est en *-y-*. On l'a comparé avec le suffixe de même forme qui sert à former des noms d'agent. Le nom d'agent peut parfois exprimer l'aspect perfectif :

finnois : *kuoliya* « mort »
estonien : *kōl'ya* « cadavre »
vepse : *kōli* « cadavre »

à côté de fi. :

luoya « créateur »
opettaya « professeur ».

En regard de ces noms d'agents, on a les prétérits :

fi. : *asu-i-n* « j'habitai »
asu-i-t « tu habitas »
asu-i « il habita »

mordve : (de *p'el'e-* « avoir peur » *kulo-* « mourir »)
p'el'-in' « j'ai eu peur »
p'el'-i-t' « tu as eu peur »

kul-i-n'ek' « nous mourûmes »
kul-i-d'e « vous mourûtes »

Quant aux différences modales, la langue primitive savait distinguer entre indicatif et subjonctif. L'indicatif n'est marqué par aucun suffixe spécial indiquant le mode. Pour ce qui est de l'impératif, il n'était autre chose à l'origine qu'un thème de présent prononcé sur un ton impérieux.

mordve : *er'a-* « vivre », *rama* « acheter », *vano-* « voir » :
er'ak « vis »
ramak « achète »
vanok « vois »

Le *-k* n'est autre que le suffixe de présent dont il a été parlé plus haut.

Quant au subjonctif, il est marqué par un suffixe spécial immédiatement lié au thème verbal et suivi du suffixe soit de présent, soit de prétérit, selon que l'on a un présent ou un prétérit de subjonctif. Le suffixe spécial du mode subjonctif était en *-n-*. Ce subjonctif servait à exprimer l'action voulue, conditionnée ou désirée. Selon les langues, il a donné un mode ou conditionnel ou optatif ou potentiel :

hongrois *men-* « aller », *mennē-k* « je voudrais aller, j'irai » ;
finnois *anta-* « donner » *anta-ne-t* « tu peux donner » ;
tchérémisse *kol-* « mourir », *kol-ne-m* « je voudrais, je désirerais mourir ».

A l'origine cette forme n'est autre chose qu'un thème de présent ou de prétérit d'un verbe dérivé. En effet le suffixe de subjonctif *-n-* coïncide avec un suffixe de dérivation déverbative. Ce suffixe exprimait tantôt l'inchoatif, tantôt l'itératif et de là s'est développé le sens d'action voulue, désirée, etc.

Une des caractéristiques des langues finno-ougriennes comme des langues samoyèdes, c'est l'extraordinaire multiplicité des suffixes de dérivation servant à former de nouveaux verbes soit à partir de verbes, soit à partir de noms. Chaque idiome possède une foule de dérivations pour exprimer les nuances fréquentative,

momentanée, itérative, causative, réfléchie et même passive. La valeur de chacun d'eux varie de parler à parler, et plusieurs se combinent volontiers ensemble.

Ainsi le verbe fi. : *yuokse-* signifie « courir », *yuokse-le-* signifie « courir çà et là ».

De façon analogue on a :

hongrois : *döf-* « cogner », *döföl-* « cogner de façon répétée » ;

ostiak : *xòł-* « entendre », *xòłənt-* « entendre de façon soutenue, écouter »

finnois : *anta-* « donner », *antautu-* « se donner, s'abandonner »

ostiak : *kur-* « prier » et *kuris'k-* « prier pendant longtemps, faire sa prière ».

Ce suffixe *-sk-* est combiné en finnois avec le suffixe *-l-* cité plus haut pour fournir des verbes comme :

uiskele- « nager çà et là » (*ui-* « nager »), etc.

Les verbes passifs ne sont pas autre chose que d'anciens verbes réfléchis dérivés dont le suffixe de dérivation a perdu peu à peu son sens premier.

A côté des formes de la conjugaison proprement dite et se mêlant à celles-ci, il existe un grand nombre de substantifs verbaux. Ils désignent l'action abstraite ou celui qui en est l'acteur ou encore le résultat de l'action. La combinaison de ces substantifs verbaux, leur emploi aux différents cas de la déclinaison et l'addition de suffixes possessifs leur permettent de jouer un rôle capital dans la phrase finno-ougrienne. Ils y tiennent lieu de phrases subordonnées introduites par des conjonctions ou des relatifs. Ils correspondent assez souvent, notamment en finnois, à nos emplois de l'infinitif précédé d'une préposition.

Ainsi « il fut empêché de partir » se dit en finnois *hän estüi lextemästä* (il se trouva empêché du fait de partir) ; « il s'instruit en lisant » se dit *lukemalla oppī* (il apprend avec le fait de lire) ; « il est venu pendant que je t'attendais » se dit *hän tuli sinua odottamassani* (il est venu au milieu de mon fait de t'attendre).

Lextämä signifie : le fait de partir (*lextä-* « idée de partir » et *mä* suffixe déverbatif), de même *lukema* « le fait de lire »

(*luke-* « idée de lire » et *-ma* même suffixe que précédemment). *Odattama* construit de la même façon sur *odotta-* « idée d'attendre ». Une forme comme *lextämästä* est le substantif verbal *lextämä* mis au cas *élatif* (correspondant à notre locution prépositive « hors de »). *Lukemalla* est le cas adessif de *lukema*. Quant à *odottamassa-ni* c'est le substantif verbal *odattama* employé à l'*inessif* et auquel a été ajouté le suffixe possessif *-ni* « mon ».

Le même substantif verbal en *-ma*, mis à un autre cas indique le début de l'action : *menen uimān* (*menen* « je vais » *ui-* « nager ») « je m'en vais nager » (je vais entrer dans le fait de nager).

Les phrases citées sont toutes finnoises. Le hongrois en effet emploie de moins en moins ce genre de construction qu'il remplace par la construction subordonnée, avec une forme personnelle du verbe.

A côté de ces substantifs verbaux, il y a une riche série de participes. Leur usage est le même, ou à peu près, que celui des participes des anciennes langues indo-européennes. Cependant le finnois leur accole volontiers un suffixe possessif, tout comme à de véritables substantifs :

hän on olevinansa viisas « il se croit sage » (il est avec son fait d'être sage), (*hän* « il », *on* « est », *olevina-* « avec le fait d'être », *-nsa* « son » suffixe possessif de la troisième pers. sing.). On a vu plus haut (page 171) *tul-tu-a-ni* « après que je suis arrivé » (après mon fait d'être venu).

Les exemples qui précèdent ont sans doute déjà fait supposer que les langues ouraliennes ignorent les préfixes. Pour rendre les mêmes idées que le français *in-* dans les adjectifs négatifs, elles recourent à un suffixe négatif :

finnois *voima* « force », *voimaton* « impuissant » ;
hongrois *hâson* « utilité », *hâsontâlân* « inutile ».

Structure de la phrase.

L'ordre des mots dans la phrase est assez libre, mais il n'a cependant pas la liberté des langues indo-européennes anciennes.

L'adjectif doit toujours précéder le substantif qu'il qualifie. De même l'absence de désinence de génitif caractéristique oblige à ne pas séparer le déterminant du déterminé et ainsi de suite.

Deux choses caractérisent la phrase finno-ougrienne : la négation et le jeu des particules de renforcement.

La négation est exprimée par un verbe négatif : ainsi le finnois dit :

en tule' « je ne viens pas »
et tule' « tu ne viens pas »
ei tule' « il ne vient pas »

Le mot *tule'* est le thème verbal de présent, sans aucun suffixe. Les formes *en*, *et*, *ei* sont la 1re, 2e et 3e personne du verbe négatif. Cela revient à dire : je dénie le fait de venir, tu dénies le fait de venir, il dénie le fait de venir.

De même, pour l'impératif négatif on se sert d'un verbe d'interdiction qui est suivi du thème nu du verbe dont on défend l'action : *älä' mene'* « ne va pas » : c'est *älä'* qui est le verbe d'interdiction (défendu, le fait d'aller). Seules les langues ougriennes ont renoncé à ce verbe négatif et au verbe d'interdiction, elles emploient des particules tout comme le latin. En revanche les langues samoyèdes les ont conservés.

Quant aux particules de renforcement, elles jouent encore dans une langue comme le finnois moderne un rôle capital. On les ajoute aux interrogatifs dès que l'on veut insister ; on les suffixe aux mots importants de la phrase : substantif, verbe, pronom. Elles renforcent l'affirmation, et également la négation. Ceci est ancien. Ainsi le *-k* de la 1re personne du singulier de la conjugaison subjective hongroise : *īrok* « j'écris » n'est que le reste d'une ancienne particule de renforcement en *-k*. : ostiak : *manlem-kə* : « oui, je vais » ; *ollem-kə* « oui, je suis ». (Voir Simonyi : *Die ungarische Sprache*, p. 350.)

Les langues de civilisation.

C'est seulement très tard que des langues de civilisation se sont formées parmi les langues finno-ougriennes. Quant aux langues samoyèdes, les conditions naturelles ont empêché les populations qui les parlaient d'atteindre même au degré le plus humble de la culture. Sans les recherches des savants occidentaux comme *Castrén*, *Kai Donner* et *Lehtisalo*, elles seraient condamnées à disparaître sans laisser le moindre vestige.

Sans culture originale, sans organisation sociale très forte, dépourvus de ces grandes aristocraties ambitieuses qui disciplinèrent les peuples indo-européens et les lancèrent à la conquête de nouvelles terres et de nouveaux progrès, les peuples finno-ougriens ont longtemps mené une vie obscure, toute végétative. Des accidents historiques qui les ont placés dès la plus haute antiquité dans le voisinage des peuples indo-européens les ont soumis par là-même à l'influence civilisatrice de ces derniers. Le vocabulaire finno-ougrien reflète encore aujourd'hui cette action qui s'est exercée en un temps très reculé.

Les mots qui servent à désigner : le grain (fi. *yüvä*), l'abeille (fi. *mehiläinen*), le miel (fi. *met-*), l'orphelin (fi. *orpo*), le cousin (fi. *orpana*), sont empruntés au groupe indo-iranien des langues de la famille indo-européenne. A ces mots s'ajoute le nom de nombre 100 : fi. *sata*, hongr. *sāz*, dans lequel on ne saurait méconnaître un emprunt à l'indo-iranien.

Plus tard, une fois séparées du tronc finno-ougrien commun, les langues ougriennes (hongrois, vogoule, ostiak) restèrent en contact avec des langues indo-iraniennes. Le nom de l'or : ostiak, *sorn'a*, *sorni*, vogoule : *suren'* rappelle le vieux perse *zaranu*, *zaranya*. De même, les finno-ougriens restés sur la Volga subissaient encore un certain temps l'influence iranienne, sans qu'on sache avec précision comment. Peut-être par l'intermédiaire des *Scythes* dont le parler était un des dialectes iraniens du Nord.

Parvenues dans la région des grands lacs russes, les tribus finnoises, caréliennes, vepses, estoniennes et lives ne tardèrent pas

à tomber sous l'influence des Baltes de langue indo-européenne et de culture plus avancée. De nombreux emprunts perpétuent de nos jours cette influence balte et sont à dater d'une époque où toutes ces tribus du groupe finnois parlaient une langue que les diversités dialectales n'avaient pas encore trop différenciée. Exemples : le nom du pâtre (fi. *paimen*), du bélier (fi. *oinas*), du foin (fi. *heinä*), du parc à bétail (fi. *tarha*), de la semence (fi. *siemen*), du mur (fi. *seinä*), du pont ou de la passerelle (fi. *silta*), de la poutre qui soutient le toit (fi. *malta*), de la hache (fi. *kives*), de la roue (fi. *ratas*), du navire (fi. *laiva*), de la voile (fi. *purye'*), de la fiancée (fi. *morsian*), de la tribu (fi. *heimo*), du ciel (fi. *taivas*), du démon (fi. *perkele'*), enfin le nom du nombre mille (fi. *tuhat*).

L'influence des peuples germaniques vint ensuite et elle n'a pas cessé de s'exercer depuis lors. Il est assez difficile de toujours distinguer entre emprunts de la langue proto-finnoise et emprunts parallèles de chacune des langues du groupe.

Parmi les plus anciens, dont le nombre est très considérable, on signalera les noms du mouton (fi. *lammas*), du fromage (fi. *yūsto*), du champ (fi. *pelto*), de la charrue (fi. *aura*), du pain (fi. *leipä*), du seigle (fi. *ruis*), du plancher (fi. *lattia*), de la chemise (fi. *paita*), de la robe (fi. *hame'*), du commerce (fi. *kauppa*), du fer (fi. *rauta*), de l'or (fi. *kulta*), de l'étain (fi. *tina*), du plomb (fi. *lüiyä*). Les termes d'organisation sociale et administrative sont également germaniques : le roi (fi. *kuningas*), le prince (fi. *ruhtinas*), la gloire ou l'honneur (fi. *kunnia*), l'état, le pouvoir (fi. *valta*), gouverner (fi. *hallita'*), juger (fi. *tuomita'*), l'emprunt (fi. *laina*), l'amende (fi. *sakko*), la redevance (fi. *vuokra*).

Enfin, les Slaves fournirent aussi un certain nombre de mots au vocabulaire des langues du groupe finnois, quoique à une époque bien plus tardive. Le mot fenêtre (fi. *akkuna*) est indiscutablement d'origine slave, ainsi que le nom du marché (fi. *turku*), de la marchandise (fi. *tavara*). Les mots du christianisme sont empruntés aux Slaves : fi. *risti* « croix », *pappi* « prêtre », *pakana* « païen », *rāmattu* « bible ». Ces emprunts datent, pour ce qui est des plus anciens, du premier millénaire après Jésus-Christ.

De même le hongrois. Sorti du voisinage des Ossètes auxquels

il doit : *gâzdâg* « riche » (ossète : *xazdug*), *kârd*, couteau (ossète : *kard*), etc., il a reçu ensuite des mots turcs, et d'abord sous une forme tchouvache qui les fait remonter à l'époque où les Hongrois parcouraient encore les steppes de la Russie du Sud. Ce sont les noms de l'orge : *ārpå*, du froment : *būzå*, de la faux : *šārlō*, le verbe moudre : *ōröl*, les noms des fruits, les termes d'élevage (*bårom* « bétail », *ökör* « bœuf », *ürū* « mouton », etc.), sans compter les noms d'instruments, des parties du vêtement. L'influence turque ne devait point s'arrêter là. Si les Turcs du Kouman et les Petchenègues n'ont laissé que des traces insignifiantes dans le vocabulaire hongrois, en revanche, le désastre de Mohács (1526) qui livra la Hongrie aux conquérants osmanlis a augmenté de beaucoup le nombre d'éléments turcs que l'on retrouve en hongrois.

Auparavant, l'établissement des Hongrois ou Magyars (prononcer *mâd'âr*) sur le territoire de la Hongrie actuelle leur avait donné pour voisins les Slaves du Sud. Beaucoup d'esclaves slaves étaient venus accroître la population magyare. Ils avaient apporté avec eux les mots slaves de civilisation. C'est de cette époque que datent les désignations de la plaine : *pustå*, du ruisseau : *påtåk*, du soc de charrue : *tšorosyå*, du lin : *len*. De même tous les termes servant à dénommer les parties de la maison, de l'ameublement, les termes des métiers (*kovātš* « forgeron », *tåkātš* « tisserand ») s'introduisent en grand nombre ainsi que le vocabulaire du christianisme : *kerestēn'* « Christ », *poyån'* « païen », *kerest* « croix », *sent* « saint », *pokol* « enfer ». Il faut y ajouter les noms des jours de la semaine : *sombåt* « samedi », *serdå* « mercredi », *tšutörtök* « jeudi », *pēntek* « vendredi ». C'est encore aux Slaves que les Hongrois doivent une partie de la terminologie administrative *kirāy* « roi », *tšāsār* « empereur », *dēsmå* « dixième » (de redevance), etc.

L'italien a succédé au slave. Le prestige de Venise, son commerce et, plus près de nos jours, l'occupation de la Lombardie par les troupes hongroises, ont entraîné l'adoption de beaucoup de termes italiens, sans parler du latin qui a été, on le sait, la langue officielle de la Hongrie presque jusqu'à notre époque.

On n'insistera pas non plus sur les multiples influences russes, tatares et autres subies par les autres langues du groupe et même par les langues samoyèdes. Il faut y ajouter toutes sortes d'influences réciproques. Ainsi le lapon ne s'est pas contenté d'emprunter un vocabulaire de civilisation au germanique et, plus spécialement à l'époque moderne, au scandinave. Dès le moyen âge, il s'est fourni chez les Finnois pour nombre de termes qui lui faisaient défaut.

Les langues de civilisation finno-ougriennes sont : le finnois, l'estonien, les langues littéraires laponnes et le hongrois.

On a vu plus haut quelle influence le groupe finnois tout entier avait subie de la part des peuples baltes, germaniques et slaves. Cette triple action s'est poursuivie par la suite.

Conquise au moyen âge par les rois de Suède, la Finlande se vit bientôt colonisée par de nombreux immigrants suédois : paysans et marchands, auxquels vinrent s'ajouter dans la suite des fonctionnaires et des administrateurs. La langue officielle étant le suédois, on imagine aisément l'infiltration qui s'est produite au cours des siècles, et aujourd'hui on aurait peine à se faire comprendre en finnois sans employer à tout instant des mots nordiques et particulièrement des mots suédois plus ou moins assimilés. La Réforme fit publier la Bible en langue populaire, c'est-à-dire en finnois, afin de la mettre à la portée de tout le monde (1548). Mais à côté de cette nouvelle tradition écrite qui allait désormais se développer de plus en plus, subsistait parmi les paysans une puissante tradition orale. Les chants populaires, recueillis au XIXe siècle par *Elias Lönnrot*, en restent l'imposant témoignage. Au cours du XIXe siècle, la langue écrite achève vraiment de se former. Des écrivains originaux, habitués à la pensée européenne et à toutes ses nuances essayèrent de s'exprimer en finnois. Ils réussirent à se forger un instrument d'expression d'une rare précision et d'une prodigieuse richesse de vocabulaire.

Une évolution parallèle se produisait pour l'estonien. Le Nouveau Testament, publié en 1632 pour la première fois, marque le début d'une littérature en langue estonienne. Soumise aux barons

allemands, l'Estonie prit à l'allemand ce que la Finlande recevait du suédois : tous les mots d'administration et de civilisation moderne. C'est encore aujourd'hui l'une des grandes difficultés qui empêchent un Finnois de lire facilement un journal estonien ou inversement.

Le lapon a vu se former en Norvège et en Suède des langues écrites, dans lesquelles sont rédigés les textes sacrés et quelques chroniques. Ces langues sont le résultat de tentatives faites pour normaliser quelques dialectes un peu plus répandus que les autres. Elles n'ont pas réussi à s'imposer au delà des localités où elles sont confinées.

Quant au hongrois, écrit dès le XIVe siècle, il a produit une littérature considérable, surtout à partir du XVIIIe siècle. La presse, le roman, les langues techniques ont développé tous les aspects de la langue, lui ont donné une souplesse, une variété d'expression au moins égales à celles de toute langue indo-européenne de civilisation.

Les termes de la technique moderne sont passés dans le finnois, l'estonien, le hongrois tout comme en allemand, en russe, en anglais. Souvent ils ont servi à former des mots en tous points comparables à ceux calqués par l'allemand ou le slave sur des modèles gréco-latins.

Au cours du XIXe siècle, le livre et la pensée allemands ont agi de plus en plus profondément sur le hongrois, sur l'estonien et par le suédois et, à côté de celui-ci, sur le finnois. On retrouve aisément dans un journal finnois, estonien ou hongrois, les formules que l'on a lues dans les journaux allemands. Seuls les mots sont d'une autre origine, mais leur assemblage est presque le même. Bien plus, des particularités frappantes de l'allemand se retrouvent en estonien. Le verbe de la proposition subordonnée s'y place en effet en fin de phrase.

La pensée germanique, allemande ou suédoise, a été prise pour modèle. A essayer de la rendre ou de l'imiter, on s'est ingénié à imiter la phrase germanique. Le finnois a développé un parfait comparable au parfait suédois :

minä olen tullut
je suis venu
(ich bin gekommen.)

L'estonien a un verbe *sāma* dont bien des emplois correspondaient à ceux du verbe allemand *werden*. L'estonien a généralisé le verbe *sāma* partout où l'allemand avait *werden*. Il a créé ainsi un futur qui était autrefois inconnu aux langues finno-ougriennes et que le finnois, son proche parent et voisin, ignore encore :

ta sāB tulema
il va venir
(er wird kommen.)

Le passif finnois, tout comme le passif des autres langues finno-ougriennes, est issu d'un verbe dérivé réfléchi. Sous l'influence de l'allemand, l'estonien s'est créé de toutes pièces une conjugaison passive où *sāma* joue le rôle de *werden*.

sē sāB tehtuD
ceci est fait
(es wird getan.)

Si le finnois et le hongrois n'ont pas été aussi loin que l'ēstonien, ils ont cependant marché dans le même sens. On reconnaît dans les propositions subordonnées du finnois les conjonctions nordiques qui les introduisent. Le mécanisme des relatifs est entièrement calqué sur celui du suédois.

Le verbe perd ses aspects. Ses formes prennent une valeur temporelle. Les suffixes de dérivation se vident de leur signification première. Ils ne sont plus que des procédés de formation de nouveaux verbes. Le lapon, le hongrois ne connaissent plus la foule des substantifs verbaux que le finnois maintient encore et qui répondent aux divers emplois de notre infinitif, de notre participe et de nos propositions subordonnées. A la simple juxtaposition de phrases courtes, succède de plus en plus la subordination, si étrangère aux langues finno-ougriennes primitives.

L'évolution phonétique tend à effacer certaines correspondances très claires. Le nominatif finnois *käsi* (main) diffère désormais dans sa forme du génitif *käden*, du partitif *kättä*, de l'illatif *käteen* (« dans la main » avec mouvement). Il a suffi que

*-*ti* passe à *-*si* à la finale et que le jeu de l'alternance *t/d* soit moins nettement senti pour obscurcir tout aussitôt un paradigme autrefois clair (**kät-i, kät-tä, käd-en, kät-ehen*). Ceci est encore plus vrai dans les formes hongroises : nominatif *kö* « pierre » et aux différents cas de la déclinaison : *köve, köv-*. De même : *yö* et *yåvå-, yåv-*.

Certains cas se confondent. La désinence *-m* étant devenue *-n* en finnois, le génitif et l'accusatif se trouvent avoir même désinence. Il n'en était pas ainsi autrefois dans les langues finno-ougriennes. Plus on va, plus chaque forme tend à se distinguer des autres, à prendre un aspect unique, particulier. On n'est plus très loin du système de la flexion des langues indo-européennes anciennes.

De pareilles innovations révolutionnent la structure linguistique finno-ougrienne. A quelques expressions ou tournures près, le lecteur européen ne se trouve pas dépaysé devant une page d'un journal finnois, hongrois et surtout estonien. La syntaxe ne lui oppose rien qui lui paraisse foncièrement original. Seul le vocabulaire lui reste étranger. Conservant leurs éléments primitifs, les langues finno-ougriennes de civilisation ont changé totalement de caractère sous l'influence des langues indo-européennes de civilisation. Elles n'en seront bientôt plus que des décalques dont seuls les matériaux rappelleront leur origine et leur continuité historiques.

A. SAUVAGEOT.

BIBLIOGRAPHIE

I. Sur les peuples de langue finno-ougrienne :

ITKONEN : *Suomensukuiset kansat* (Les peuples finno-ougriens). Helsingfors, 1921.

II. Sur les langues finno-ougriennes :

J. SZINNYEI : *Finnisch-ugrische Sprachwissenschaft* (Collection Göschen. 2e édit., 1922).

SIMONYI : *Die ungarische Sprache* (Strasbourg. Trübner, 1907).

III. Sur les langues samoyèdes :

Kai DONNER : *Ulber die anlautenden labialen Spiranten und Verschlusslaute im Samojedischen und Uralischen* (Helsingfors, 1920. *Mémoires de la Société finno-ougrienne*, XLIX).

La bibliographie détaillée se trouve dans les ouvrages indiqués ci-dessus. La plupart des études récentes parues ont été l'objet de recensions dans :

Journal de la Société finno-ougrienne de Helsingfors (I à XXXIII. Helsingfors).

Mémoires de la Société finno-ougrienne (I à XLIX. Helsingfors). Les articles sont rédigés en français, en allemand, en anglais et en finnois.

Finnisch-ugrische Forschungen (Revue de linguistique et d'ethnographie finno-ougriennes). I à XV Helsingfors. Les *Anzeiger* parus en appendice fournissent toute la bibliographie concernant le domaine finno-ougrien jusqu'en 1906. La revue est rédigée en allemand.

LANGUES TURQUES, LANGUES MONGOLES

ET

LANGUES TONGOUZES

GÉNÉRALITÉS

On a souvent supposé que les langues turques, mongoles et tongouzes constituaient le rameau « altaïque » d'une grande famille « ouralo-altaïque » ou « touranienne ».

Cette théorie dont on a, à tort d'ailleurs, attribué la paternité à Abel Rémusat, est née vers 1830 et bien que s'étant heurtée dès le début à l'opposition d'un savant comme Otto Böhtlingk, a fini par prendre la valeur d'un dogme chez beaucoup d'auteurs.

Cependant depuis les progrès réalisés par la linguistique finno-ougrienne, on a cessé de grouper le turc avec le finno-ougrien.

Le turc, le mongol et le tongouze présentent entre eux des ressemblances frappantes.

Il faut reconnaître cependant que les arguments invoqués par les partisans de l'unité d'origine de ces trois groupes sont généralement du domaine du vocabulaire (grand nombre de mots communs) ou du domaine de la phonétique (harmonie vocalique), ou enfin de celui de la syntaxe. Or ce sont là des faits qui peuvent s'expliquer soit par l'emprunt (voir l'étude de Vladimirtsov sur les mots turcs passés en mongol, *Zapiski vost. otd. imp. roussk. arxeol. ob.*, t. XX, 1911), soit par l'influence qu'exerce une langue sur une autre. Il s'agit, en effet, de peuples qui sont non seulement voisins, mais qui ont été mêlés et brassés par une série ininterrompue de guerres et de migrations. Si l'on pouvait se contenter de considérations tirées de la syntaxe, il faudrait donner raison à ceux qui ont rangé dans la même famille le japonais et le turc.

Les ressemblances d'ordre morphologique, qui seraient les plus

concluantes, se réduisent encore à assez peu de chose : similitude dans la structure syllabique des mots et dans le mécanisme des variations grammaticales (suffixation). Quant aux morphèmes, ceux qui se laissent ramener à l'unité sont peu nombreux ; on en verra cependant des exemples nets : ce sont, notamment, le suffixe du locatif qui est *-dä* dans ces trois groupes de langues, le suffixe *-ki* qui accompagne le même locatif (p. 228) et le suffixe du génitif. Les formes du pronom personnel se ressemblent beaucoup (Friedrich Müller, *Das Personal-Pronomen der altaischen Sprachen, Sitz. Kais. Ak. der Wiss. in Wien*. 1895).

C'est en multipliant ces rapprochements et en en découvrant de nouveaux qu'on pourra, espérons-le, établir rigoureusement la parenté des langues turques, mongoles et tongouzes.

Les cinq faits suivants viennent compliquer l'étude de la grammaire comparée de ces langues:

1° Obscurité des origines des peuples qui les parlent. Dès le deuxième millénaire avant notre ère, les Chinois se trouvaient en contact avec des peuplades « barbares », au Nord du Fleuve Jaune dans les provinces de Chan-si, Chàn-si et Kan-sou. Le nom du principal de ces peuples change d'orthographe au cours des siècles dans les chroniques chinoises, mais ses diverses formes, dans tout le premier millénaire avant l'ère chrétienne, semblent bien représenter déjà le nom même des Hiong-nou, dont l'énorme empire s'étendait au Nord de la Grande Muraille construite contre eux à partir de 214 av. J.-C. Cet empire fut détruit au IIIe siècle de notre ère. Deguignes, dès 1756, proposa d'identifier les Hiong-nou avec les Huns d'Attila et cette hypothèse, sans être vraiment prouvée, reste très vraisemblable (voir les travaux de Hirth et de Franke). Mais qui étaient les Huns? Des Turcs, comme on l'a cru le plus généralement depuis Klaproth et comme on vient de l'affirmer à nouveau en Allemagne? des Mongols? ou des Tongouzes (voir Shiratori, *J. As.*, 1923)? On l'ignore encore. Ce n'est qu'à partir de la constitution de l'Empire des T'ou-Kiue (Turcs) au VIe siècle (voir plus loin, p. 194) que l'on commence à voir plus clair dans l'histoire de l'Asie Centrale.

2° Extrême mobilité des peuples. Il suffit de rappeler, dans cet ordre d'idées, les invasions des peuples qui traversèrent les steppes de la Russie méridionale et pénétrèrent jusque dans les Balkans : Huns, Avars (probablement des Mongols), hordes turques telles que Petchénègues et Oghouz (IXe siècle), Comans ou *Polovtsị* (XIe siècle), invasion mongole du XIIIe siècle, sans compter celles des Seldjoukides et des Ottomans.

Citons encore des migrations mongoles relativement récentes : dans le courant du XVIIe siècle, les Torgoutes obtinrent du gouvernement russe la permission de s'installer entre la Volga et le Don et en 1771 une partie de cette population revint, au nombre de 30 000 familles, et au prix des pires souffrances, se soumettre au gouvernement chinois ; en 1688, mille familles khalkha (Mongols du Nord) pressées par les Oïrat (fédération de Mongols Occidentaux) traversèrent la frontière russe ; ils sont connus actuellement sous le nom de Bouriats de la Selenga.

Il y a aussi un mouvement de migration de peuplades tongouzes de Russie en Mandchourie.

3° Changement de langue. On peut relever le caractère éphémère de certains dialectes ou langues, conséquence probable de la mobilité des peuples et de l'absence d'une civilisation nationale jouissant d'un grand prestige. C'est ainsi qu'à en croire certains textes chinois, les Kirghizs — dont la présence entre l'Ob et l'Iénisseï est signalée dès le Ier siècle av. J.-C. — n'étaient pas à l'origine un peuple de langue turque (Radloff, *Alt. Inschr.*, p. 425). Les *Mišer* (les Mechtcheriaks des Russes) et les *Tepter*, peuplades d'origine diverse, n'ont été turquisés que postérieurement. Les Tatars de l'Iénisséï sont, en grande partie, des descendants de peuplades samoyèdes.

L'exemple suivant, presque contemporain, est assez déconcertant : les Kamassi, peuplade de langue samoyède, commencent à parler turc (dialecte de la Katcha) en 1840 ; en 1860, ils ont oublié leur ancien parler, mais en 1890, ils abandonnent le turc pour le russe (Katanof, *Ouriankhaï*, pp. XVII-XVIII). Deux changements de langue en cinquante ans !

4° Similitude relativement grande entre les dialectes ou

langues de même groupe, surtout dans les groupes turc et mongol. Dans le groupe turc, seuls le yakout et le tchouvache offrent un aspect aberrant. Il en résulte que les formes ou les mots du turc commun qu'on peut restituer ressemblent étrangement aux formes ou aux mots des différents dialectes actuellement parlés.

5° Lenteur de l'évolution des différents dialectes. Ce fait est solidaire du précédent. Les documents les plus anciens (VIIIe siècle pour le turc, XIIIe siècle pour le mongol) nous livrent des langues qui diffèrent relativement très peu des parlers modernes qui leur ont succédé.

Traits communs aux langues turques mongoles et tongouzes

Nous énumérons ici les principaux traits communs aux trois groupes de langues, non pour fournir des arguments à la théorie de l'unité de leur origine, mais pour éviter d'inutiles redites. Les ressemblances entre le turc et le mongol sont plus frappantes qu'entre le tongouze et les mêmes langues.

I. — Phonétique.

1° « Harmonie (assimilation) vocalique », phénomène dont il a déjà été question à propos des langues finno-ougriennes et qui se manifeste à des degrés de développement variable dans les différents dialectes (pour plus de détail, voir p. 204).

2° Répercussion de l'harmonie vocalique sur certaines consonnes, surtout les gutturales. Les gutturales prépalatales *k'* (*x'*) et *g'* ne se rencontrent que dans le voisinage des voyelles antérieures, les gutturales postpalatales *q* (*x*) et *g* (*g*) dans celui des voyelles postérieures seulement. La gutturale d'un suffixe prend l'une ou l'autre forme, suivant qu'il se joint à un mot dont la dernière voyelle est de la classe antérieure ou postérieure et les trois groupes de langues ont des signes différents pour ces deux séries de consonnes, quel que soit l'alphabet en usage. Quant aux autres consonnes dont il existe des doublets analogues — l'un antérieur (palatal), l'autre postérieur (guttural ou vélaire) —

comme par exemple en mandchou et osmanli pour *t* et *d*, en osmanli moderne pour *s*, en turc de l'Orkhon pour un grand nombre de consonnes (voir p. 216), il resterait à examiner s'il ne s'agit pas simplement de graphies destinées à l'origine à suppléer à l'insuffisance de la notation des voyelles. Certains dialectes turcs distinguent deux sortes d'*l* : *ł* vélaire ou creux dans les mots de la classe postérieure et *l* ordinaire ou plat dans les mots de la classe antérieure, mais sans les discriminer dans l'écriture.

3° Tendance à éviter une consonne continue sonore à l'initiale du mot, tendance sensible surtout en turc où l'on ne trouve dans cette position aucune des consonnes suivantes : *v* (sauf en osmanli), *n* (sauf dans le mot *näñ* ou *nä* « chose, rien »), *ñ*, *z*, *d*, *g* (postpalatal), *r*, *l*. Le son *m* initial y est tardif (voir p. 202). La consonne *ñ* initiale est inconnue aussi au mongol et la liquide *r* initiale au mongol et au tongouze.

4° Rôle très réduit des semi-voyelles (le son *w* n'apparaît que comme tardif et rarement). Le son *y* est employé en valeur de consonne ordinaire.

5° Importance des voyelles. Les voyelles, généralement d'un timbre net (peu de voyelles mixtes), sont des éléments en général très stables.

6° Instabilité, plus ou moins accentuée, de l'*n* final.

II. — Structure syllabique des mots.

Les consonnes ne sont jamais géminées (dans les racines ou dans les suffixes pris isolément).

Jamais un mot ne commence par un groupe de consonnes.

Les syllabes sont rarement fermées par deux consonnes, à moins que la première ne soit une liquide (*r*, *l*). En mandchou surtout, emploi fréquent des syllabes ouvertes (presque autant qu'en japonais).

III. — Morphologie.

1° Pas de genre grammatical.

2° Deux nombres seulement (pas de duel).

3° Emploi de racines « nues ». En principe, toute racine peut être employée dans la phrase sans le secours d'aucun suffixe, c'est-à-dire comme mot non fléchi ou comme *base* (si l'on adopte ce terme commode pour désigner les mots non fléchis). Les bases sont *verbales* (exprimant une action, un devenir ou un état) ou *nominales* (désignant un objet).

Toute base verbale a la valeur d'un impératif sg., 2e personne.

Toute base nominale a, en principe, la valeur d'un sujet au singulier.

4° Les variations grammaticales s'obtiennent surtout au moyen de suffixes (pas de préfixes, ni d'infixes). Les suffixes sont faciles à reconnaître parce qu'ils n'altèrent pas d'une façon appréciable la structure phonétique du mot « nu » (base) auquel ils se joignent et parce que les changements phonétiques qu'ils subissent eux-mêmes se réduisent à peu de chose, du moins dans l'état actuel de ces langues (les principaux sont commandés par le phénomène de l'harmonie vocalique). Ces observations sont vraies surtout pour le turc. Le mongol, dans une faible mesure, et, dans une plus large, les dialectes tongouzes, offrent des exemples de flexion (par fusion plus intime entre le mot et le suffixe).

Étant donnée l'importance du rôle que jouent les suffixes en turc, mongol et tongouze, les renseignements donnés dans tout le présent chapitre sur la structure de ces langues porteront principalement sur ces morphèmes.

5° La grande masse de ces suffixes peut être ramenée, assez facilement, à deux types principaux : les suffixes de dérivation et les suffixes désinentiels.

a) Suffixes de *dérivation*. — Ils servent à former avec une base donnée une base nouvelle qu'on peut appeler *base dérivée*.

Qu'on parte d'une base nominale ou d'une base verbale, on peut former des bases dérivées aussi bien verbales que nominales. Toutes les combinaisons sont possibles.

Parmi les suffixes de dérivation, qui relèvent plutôt de la lexicographie que de la grammaire, il en est cependant qui jouent

un rôle nettement grammatical. Ce sont les suffixes exprimant la catégorie de la « voix » : passive, réfléchie, etc.

b) Suffixes *désinentiels*. — Est désinentiel tout suffixe qui sert à fléchir une base, c'est-à-dire tout suffixe qui enlève à une base verbale son caractère d'impératif de la 2ᵉ pers. du sing. et à une base nominale celui de sujet (au singulier) de la proposition.

Le schéma du mot fléchi (qui est aussi le schéma le plus complet du mot dans les langues qui nous occupent) est le suivant :

Mot fléchi = base élém. + un ou plus. suff. de dér. + un ou plus. suff. désin.

Base dérivée

Ex. : turc *tšał-* : base élém. du verbe « frapper, jouer d'un instrument à musique ».
tšał-gį : base élém. nom. dérivée « instrument à musique, musique ».
tšał-gį-džį : autre base élém. dér. « musicien ».
tšał-gį-džį-łar : mot fléchi « musiciens ».
tšał-gį-džį-łar-a : autre mot fléchi « aux musiciens ».

c) Suffixes *mixtes*. — Certains suffixes peuvent être considérés comme mixtes parce qu'ils tiennent à la fois du suffixe de dérivation et du suffixe désinentiel. Il en est ainsi de certains suffixes pronominaux (suffixe possessif par ex.) qui peuvent laisser à un nom sa qualité de sujet au singulier (comme les suffixes de dérivation), mais qui en cas de cumul avec d'autres désinences (celle du pluriel par ex.) prennent place après celles-ci :

turc *at* « cheval », *at-įm* « mon cheval », *at-łar-įm* « mes chevaux ».

Sont également mixtes les suffixes qui servent à constituer des formes nominales du verbe (noms verbaux, noms d'action, participes et gérondifs). En tant que suffixes nominaux ils sont suffixes de dérivation (puisqu'ils fournissent des bases nominales comme par ex. le suffixe du participe osmanli -(*y*)*än* dans *dik-*

än : 1° « piquant, qui pique » ; 2° « épine »), mais en tant que suffixes verbaux, ce sont des éléments désinentiels (puisqu'ils font perdre à la base verbale à laquelle ils se joignent sa valeur d'impératif : *dik !* « pique ! ».

6° L'élément initial du mot est représenté par la racine (conséquence de l'absence de préfixes).

7° Au lieu de prépositions des langues indo-européennes et sémitiques on trouve des *postpositions*.

8° Une seule conjugaison dans chaque langue (abstraction faite des changements commandés par la loi d'harmonie vocalique).

9° Grande variété des suffixes exprimant la « voix » : il existe non seulement des verbes passifs, réfléchis, mais aussi réciproques, négatifs, impossibles, etc.

Remarque. — Nous avons déjà fait allusion aux quelques suffixes qui sont les mêmes en turc, mongol et tongouze. Nous reviendrons sur ces morphèmes dans l'exposé qui suit.

IV. — Syntaxe.

1° Ordre des mots inspiré du principe suivant : *tout élément secondaire précède l'élément principal* :

a. Le déterminant se place avant le déterminé.
   α. L'épithète avant le nom.
   β. Le complément déterminatif avant le mot complété.

b. Le complément-régime (direct, indirect, de circonstance) se place avant le mot qui le régit.

c. Le sujet se place avant le prédicat.

d. Toute forme verbale prend place à la fin du groupe dont elle fait partie (conséquence des règles *b* et *c*).

2° Absence presque totale de conjonctions et de pronoms relatifs. Les propositions subordonnées ou incidentes à formes verbales personnelles des langues indo-européennes sont remplacées par des quasi-propositions terminées par des formes nominales du verbe (noms d'action, participes et gérondifs, voir p. 212).

Remarque. — Indépendamment des traits communs signalés ci-dessus, on trouvera dans le détail de l'exposé une série d'analogies de structure (morphologie et syntaxe) dont quelques-unes sont remarquables.

LANGUES TURQUES

Le nom des « Turcs ». — Comme nom commun le mot « turc » (*türk*) signifie « puissance » (F. K. W. Müller, *Uigurica*, II, p. 67) et, d'après Kachghari, « arrivé au milieu de sa course (soleil à midi), arrivé à maturité (fruit, homme) ». (Voir *Dīvānu-lugāt-it-Turk*, I, 294-295.)

Comme nom de peuple, le mot *türk* n'apparaît qu'au VI[e] siècle, au moment où sur les ruines de l'empire des Jouan-Jouan (Avars) se forme, en 552, celui des T'ou-Kiue (nom sous lequel les annalistes chinois désignaient les Turcs [1]. Les T'ou-Kiue orientaux ou septentrionaux occupaient la Mongolie, et l'habitat des T'ou-Kiue occidentaux se trouvait au Nord des T'ien-Chan depuis le lac Barkoul à l'Est jusqu'aux monts d'Alexandre à l'Ouest. Cet empire qui fut détruit au milieu du VIII[e] siècle et remplacé dans sa partie est par celui des Ouïgours, s'étendait à son apogée de la Caspienne à la Corée.

Plus tard, au XI[e] siècle, Kachghari (voir p. 203) oppose les habitants sédentaires du Turkestan chinois, de la Dzoungarie et du Sémiretchié, qu'il appelle les « Turcs », aux Oghouz et autres nomades occidentaux, mais il nomme « turque », sans restriction, la langue parlée par l'ensemble de ces peuples.

Depuis cette époque le mot *türk* a été employé par les indigènes eux-mêmes, surtout pour désigner leur langue maternelle. Comme terme ethnique ils se servirent d'autres noms, parfois de celui de leur religion (les musulmans), et le nom de *türk*

1. La transcription chinoise T'ou-kiue s'explique mieux par une forme intermédiaire *türk-üt*, pluriel mongol du mot *türk*. Il est en effet probable que les Chinois ont entendu pour la première fois ce mot de la bouche des Jouan-Jouan fugitifs (P. Pelliot, *L'origine de T'ou-kiue, nom chinois des Turcs*, T'oung-Pao, 1915).

fut en usage surtout chez les étrangers. C'est seulement depuis le réveil du sentiment national (xx[e] siècle) que l'emploi s'en généralise parmi les intéressés eux-mêmes.

Ajoutons qu'en France le mot « turc » est pris souvent dans un sens étroit pour désigner les Turcs de Turquie ou Ottomans (mieux : Osmanlis). Les auteurs qui ont voulu rendre à ce mot sa portée véritable (= « peuples de langue turque ») se sont crus obligés de souligner leur intention par un artifice orthographique, en écrivant « turk » au lieu de « turc ». Les Russes ont fait quelque chose d'analogue, en créant dans le même but à côté du mot courant *turok*, pl. *turki*, celui de *t'urk*, pl. *t'urki* (tous ces mots portent l'accent tonique sur la voyelle *u*). C'est à ces hésitations qu'il faut attribuer la création du terme de « turco-tatares » appliqué généralement à ces langues qu'il est plus simple de désigner sous le nom de « langues turques » (le nom de *tatar* désignait à l'origine un peuple de langue mongole).

Distribution géographique des langues turques. — L'aire géographique occupée par les différents dialectes turcs est relativement très vaste (voir la planche 6) : elle touche, avec des solutions de continuité assez restreintes, au 21[e] degré de longitude est, en Macédoine, et au 160 de la même longitude, dans la région des Yakoutes en Sibérie, sur le cours de la rivière Kolym. En réunissant les données statistiques fragmentaires et incertaines qu'on trouve dans différents ouvrages à celles qui sont fournies par le dernier recensement russe, de 1897, on peut admettre qu'à cette date le total des individus parlant turc s'élevait à environ 30 millions. Il est difficile de dire dans quelle proportion cette population a augmenté durant le quart de siècle qui a suivi. Une publication officielle russe assez récente (*Aziatskaïa Rossia*, Pétrograd, 1914, en deux vol. in-4° avec atlas) permet d'établir que de 1897 à 1911, soit en quatorze ans, la population turque de l'Asie Russe avait augmenté de 1.258.044 individus, ce qui suppose un accroissement annuel de 1.2 p. 100, taux plus élevé que celui indiqué pour la Sibérie seule par Patkanov qui parle d'un accroissement annuel de 0.4 à 0.7 p. 100 (*Keleti Szemle*, 1908, p. 74). En adoptant ce taux de 1.2 p. 100 pour l'augmen-

tation annuelle de toute la population de race turque, on obtiendrait pour les vingt-cinq dernières années un surplus de 9 millions d'individus, chiffre qui est inférieur à ceux dont il est parfois question à propos de l'accroissement de la population en Russie, en général (15 p. 100 pour les neuf années ayant suivi le recensement de 1897, et pour l'ensemble de l'Asie russe, 50 p. 100 en quatorze ans !)

Si l'on admet qu'il existe actuellement 39 millions d'individus parlant turc (abstraction faite des pertes occasionnées par la Grande Guerre) et si l'on réfléchit à l'étendue du territoire qu'ils occupent, on est frappé de la densité insignifiante que suppose une pareille répartition. Il est étonnant que cette dispersion n'ait pas contribué davantage à la différenciation des dialectes turcs.

Des événements plus ou moins récents (la perte des Balkans par la Turquie) ont rendu instables certaines populations turques. Elles refluent vers l'Asie dans un mouvement de migration qui ne s'est pas encore arrêté de nos jours. En 1881, il y avait 727.000 Turcs en Bulgarie. En 1910, ils n'étaient plus que 484.000 (Lamouche, *La Bulgarie*, Paris, 1923, p. 44).

On trouvera ci-dessous la liste des peuples de langue turque (voir aussi la planche 6). Les indications numériques ont été empruntées pour les Turcs de Russie au relevé général du recensement de 1897, publié par N. Troïtsky à Saint-Pétersbourg, en 1905, et pour les autres pays, au travail de M. Aristov paru dans la *Jivaia Starina* de 1896 et quelques autres ouvrages.

I. *Turcs de la Sibérie et de la Mongolie.*

1. Yakoutes et Dolganes (999) (carte : 1)...... 228.739
2. Peuplades turco-tatares de l'Altaï : Tatars de l'Altaï proprement dit (carte : 2^a) ; Tatars de l'Altaï septentrional ou Kouznetskiyés (régions de l'Iénisséï et de l'Ob supér., carte : 2^b) ; Tatars de l'Abakan (carte : 2^c).......
3. Barabas (carte : 3^a) et Tatars de la Sibérie occidentale, régions de l'Irtiche et de la Tobol (carte : 3^b)......................

(2 et 3) 210.154

4. Ouriankhaïs ou Soïotes de Russie et de Mongolie (carte : 4), environ...............		45.000
5. En Mongolie : Koktchouloutounes (carte : 4^2) et Kotones près du lac d'Oubsa (nombre inconnu).		

II. *Turcs de l'Asie centrale.*

1. Kazak-Kirghizs ou Kirghiz-Kaïssaks (carte : 5) :		
a. En Russie d'Asie........	3.881.066	
b. A Khiva et Boukhara.....	44.000	
c. En Chine..............	101.000	
TOTAL.........	4.026.066	4.026.066
d. Kiptchaks (carte : 6)................		45.353
e. Kara-Kalpaks, qu'on rattache également aux Euzbegs (carte : 7)............		111.799
2. Kara-Kirghizs ou Bouroutes (carte : 8) :		
a. En Russie..............	201.682	
b. En Boukharie...........	4.000	
c. En Chine...............	10.000	
TOTAL.........	215.682	215.682
3. Turcomans ou Turkmènes (carte : 9) :		
a. Dans l'Asie centrale russe..	248.651	
b. A Khiva et Boukhara.....	203.750	
c. En Afghanistan et Perse...	80.000	
d. Au Caucase	24.522	
e. En Russie d'Europe......	281.357	
TOTAL.........	838.280	838.280
4. Euzbegs (Ouzbegs, Ouzbeks) [carte : 10] :		
a. Dans l'Asie centrale russe..	534.825	
b. A Khiva et Boukhara.....	1.257.500	
c. En Afghanistan.........	200.000	
TOTAL.........	1.992.325	1.992.325

5. Sartes (carte : 11) :

a. Dans l'Asie centrale russe.	1.458.128	
b. En Boukharie...........	700.000	
c. En Afghanistan..........	100.000	
TOTAL..........	2.258.128	2.258.128

6. Turcs du Turkestan oriental (carte : 12) :

a. Turcs de Kouldja et du Turkestan chinois......	1.500.000	
b. Tarantchis établis dans le Turkestan russe.......	60.999	
c. Kachghariens établis dans le Turkestan russe........	41.312	
TOTAL.........	1.602.311	1.602.311

7. Turcs du Kansou (carte : 13) : Salars et Ouïgours jaunes, ces derniers, d'après Malov, au nombre de........................ 2.000
8. Tatars émigrés de la Russie d'Europe et de la Sibérie en Asie centrale russe......... 60.197
9. Turcs de l'Asie centrale dont le caractère ethnique n'a pu être précisé par le recensement russe....... .... 439.930

III. *Turcs du Sud-Ouest et de la Transcaucasie* (Caucase méridional).

1. Turcs-Osmanlis (Asie Mineure et Balkans) [carte : 14] y compris des Turcomans et des Yuruks (nomades), environ............. 10.000.000
2. Turcs-Osmanlis établis en Russie, presque tous (142.676) en Transcaucasie (carte : 14)................... .... 153.032
3. Azerbeïdjanis de la Transcaucasie (carte : 15) 1.475.553
4. Azerbeïdjanis de la Perse (carte : 15)........ 2.000.000
5. Karapapakhs (carte : 16)................ 29.902

IV. *Turcs de la Russie d'Europe et de la Roumanie.*

1. Turcs du Caucase septentrional :		
a. Karatchaïs (carte : 17)....	27.223	
b. Koumiks (carte : 18).....	83.408	
c. Kara-Nogaïs (carte : 19)..	64.080	
d. Tatars Kabarda ou des Montagnes (Gorskiyé) : Ourousbis, Balkars, Khoulams, etc. (carte : 20)...	34.232	
TOTAL	208.943	208.943
2. Turcs de l'Oural :		
a. Bachkirs (carte : 21).....	1.321.363	
b. Mechtcheriaks ou Michers (carte : 21)............	53.847	
e. Tepters (carte : 22)......	117.734	
TOTAL	1.492.944	1.492.944
3. Tatars (et Caraïtes) d'Europe :		
a. Tatars du bassin de la Volga et de la Kama (carte : 23) : gouvernements de Kazan (675.419), Oufa (184.817), Samara (165.191 dont 331 Caraïtes), Simbirsk (133.977), Viatka (125.514), Saratov (94.693), Orenbourg (92.926), Penza (58.530), Astrakhan (52.799 dont 404 Caraïtes), Perm (46.711), Nijni-Novgorod (41.339), Tambov (16.976) Riazan (5.033)...	1.693.925	
b. Tatars et Caraïtes de la Crimée et du Sud de la Russie : Crimée (196.854 dont		

5.600 Caraïtes), Prov. du Don (2.978 dont 99 caraïtes), Ekaterinoslav (17.253 dont 359 caraïtes), Odessa (3.152 dont 2.008 Caraïtes)	220.237	
c. Tatars des provinces occidentales de la Russie, de Lituanie et de Pologne : gouvernements de Kharkov (1.358 dont 255 Caraïtes), de Volhynie (3.651), Minsk (4.059), Grodno (2.420), Podolie (2.296), Kovno (1.409), Vilna (1.4.393), Pologne (4.336 dont 57 femmes seulement). Voir aussi plus bas : 4........	20.922	
d. Tatars et Caraïtes dans les deux capitales : Petrograd (5.994 dont 331 Caraïtes), Moscou (5.469 dont 347 Caraïtes) et dans les autres gouvernements russes en tout..................	19.507	
TOTAL.........	1.954.591	1.954.591
4. Lakh-Caraïtes ou Caraïtes polonais (dialecte coman) :		
à Poniewierz (gouvernement de Kovno)............	203	
à Luck (pron. *łutsk*; Volhynie)..................	166	
à Troki (gouvernement de Vilna)	576	
à Halicz (Galicie), en 1903..	192	
TOTAL	1.137	1.137

5. Tchouvaches (carte : 24)................	843.755
6. Gagaouzes de Bessarabie (christianisés) (carte : 26)..........................	55.790
7. Nogaïs de la Dobroudja ou Tchitakhs (carte : 27), 27.685, plus 12.464 Turcs, en tout, en 1895...........................	40.349
TOTAL GÉNÉRAL..	30.331.960

Voici d'autre part, pour les Turcs de l'Asie Russe quelques chiffres plus récents donnés par l'ouvrage déjà cité (*Aziatskaïa Rossia*), chiffres qu'on pourra comparer avec ceux de 1897 : Kirghizs (y compris les Kara-Kirghizs), 4.692.384 ; Sartes, 1.847.420 ; Euzbegs, 592.150 ; Turkmènes, 290.170 ; Yakoutes, 245.406 ; Karakalpaks, 134.313 ; Tarantchi, 83.000 ; Kiptchaks, 60.785 ; Kachghariens, 54.832, etc., en tout 8.363.029 Turcs, au lieu de 7.104.985 en 1897.

Presque tous sont musulmans-sunnites du rite hanéfite (en Russie : musulmans, 89.52 p. 100 ; orthodoxes, 10.05 p. 100 ; chamanistes, 0.29 p. 100 ; caraïtes, 0.07 p. 100 ; israélites, 0.04 p. 100 ; bouddhistes, 0.01 p. 100 ; grégoriens, 0.01 p. 100, et raskolniks, 0.01 p. 100).

L'islam s'est répandu non seulement aux dépens du chamanisme et du bouddhisme, mais aussi du christianisme nestorien et du manichéisme (reconnu officiellement au VIIIe siècle par les Turcs-Ouïgours).

Classification des dialectes turcs.

A défaut d'une classification vraiment scientifique des dialectes turcs, nous donnons ici, avec de légères modifications, celle de Melioranski (article dans l'Encyclopédie russe de Brockhaus et Efron) qui est elle-même un résumé de la classification de Radloff (*Phonetik der Nördlichen Türksprachen*, Leipzig, 1883, p. 280 et suiv.). Elle distingue quatre groupes de dialectes :

I. *Dialectes orientaux* : dialectes de l'Altaï proprement dit

(altaï, teléoute); dialectes de l'Altaï septentrional (dial. de la Lebed, chor); dialectes de l'Abakan proprement dit (sagaï, koïbal, dial. de la Katcha); kizil ou des Youss (les riv. *Aq-üs* et *Qara-üs*, se rattache aux dialectes de l'Abakan); baraba; kuérik (ou de la riv. Tchoulym); ouriankhaï, karagasse.

Caractéristiques. — Consonnes initiales et finales toujours sourdes (*p*, *t*, *š*, *s*, *tš*, *k* et *q*) ; consonnes intervocaliques toujours sonores (*b*, *d*, *ž*, *z*, *dž*, *g'*, *g*). Le *p* (anc. *b*) initial devient *m* lorsque la voyelle suivante est elle-même suivie d'une nasale (*m*, *n*, *ñ*). C'est dans ce groupe de dialectes que l'harmonie vocalique reçoit son application la plus large (voir p. 204).

II. *Dialectes occidentaux* : kirghiz (kazak-kirghiz, kara-kirghiz, kara-kalpak); dial. de l'Irtiche (tourali, kurdek, dial. de Tobolsk, de Tioumène) ; bachkir (b. de la plaine, b. de la montagne) ; dialectes de la Volga (micher, de la Kama, de Simbirsk, de Kazan, de Belebeï, de Kassimov).

Caractéristiques. — Présence à l'initiale des sonores *b* et (plus rarement) *d*. Passage de *š* à *s* et de *tš* à *ts*. Altération fréquente du timbre des voyelles, notamment dans le dialecte de Kazan où il existe toute une série de voyelles à timbre mixte et indéterminé.

III. *Dialectes de l'Asie centrale* : tarantchi ; de Hami ; d'Aksou ; de Kachghar; de Yarkend; sarte; dial. de Kokand; dial. des Euzbegs de la vallée de Zeravchan, dial. de Boukhara, de Khiva.

Caractéristiques. — Les voyelles *o*, *ö* apparaissent seulement à la 1re syllabe du mot. La voyelle *ị* est remplacée par un *i* indifférent au point de vue de l'harmonie vocalique. Pas de *ł*. Dans le Turkestan chinois, on trouve une harmonie vocalique régressive (au lieu d'être progressive comme ailleurs). Exemple : *bäš-i* pour *baš-i*.

IV. *Dialectes du Sud* : turcoman, azerbeïdjani (azéri); dial. du Caucase; anatolien (de Khondavendguiar, de Karaman); dialecte de Crimée; osmanli. Il faut y ajouter le gagaouze de Bessarabie.

Caractéristiques. — Les voyelles *o*, *ö* apparaissent à la 1re syllabe du mot (sauf dans certains parlers, celui d'Aïdin p. ex.). Les consonnes initiales sonores (*d*, *g* et *b*) sont plus fréquentes que dans

les autres langues turques. La gutturale continue postérieure, *g*, tombe fréquemment à l'intérieur et à la fin des mots. Conservation de l'ancien *v*, remplacé par *p* et *f* dans les autres langues. En position devant une voyelle suivie elle-même de *r* ou *z*, la consonne *b* se transforme fréquemment en *v* (*bar-maq* > *var-maq* « partir, aller », *bar* > *var* « existant », *bir-mäk* > *ver-mek* « donner » ; *yaɫbar-maq* > *yaɫvarmaq* « supplier »).

Les quatre groupes de dialectes énumérés ci-dessus sont relativement peu différenciés. Les parlers turcs se ressemblent beaucoup entre eux.

Deux dialectes cependant forment exception à cet égard, c'est le yakoute et le tchouvache. Leur isolement leur a permis d'évoluer en s'écartant du type turc commun.

La linguistique historique des langues turques est encore dans l'enfance malgré la très importante découverte de M. Vilhelm Thomsen, l'illustre savant danois qui a réussi à déchiffrer en 1893 les caractères turcs runiformes. Les inscriptions runiformes étaient connues déjà en 1730 (Pallas) ; signalées plus récemment au monde savant par Yadrintsev (1889), elles ont été relevées par Heikel et donnent les textes turcs datés les plus anciens (734 ou 735 de J.-C.). Les plus importants de ces monuments (des stèles) se trouvent près de la rivière Orkhon, affluent de la Selenga (entre le lac de *Košo-Tsaydam* et le *Kökšün-Orxon*), à environ 60 kilomètres du couvent d'*Erdeni-dzao* qui occupe lui-même l'emplacement de l'ancienne capitale mongole de Karakoroum.

Le vocabulaire turco-arabe écrit vers 1074 à Kachghar par Mahmoud Kachghari nous montre que les dialectes turcs se divisaient, au XIe siècle de notre ère, en deux groupes : l'un occidental, comprenant l'oghouz, le petchénègue, le kiptchak, le turc-bulgare, et l'autre oriental, comprenant l'ouïgour, le kachgharien et d'autres tels que le *tšigil*, le *yagma*, le *yapagu*, etc.

Or le même ouvrage permet d'établir une filiation directe entre l'oghouz du XIe siècle et l'osmanli actuel.

Il en résulte cette constatation très importante au point de vue de l'histoire de la langue turque que dès le XIe siècle la dispo-

sition des dialectes turcs correspondait en gros à leur situation actuelle, ce qui implique la réhabilitation, tout au moins partielle, du vieux terme de turc « oriental » employé par opposition à turc osmanli.

L'ouvrage de Kachghari permet de constater qu'à la même époque le son interdental *đ* (disparu depuis en se transformant en *y*, *d*, *t* ou *z*) était encore largement employé, ainsi que la labiale *v* qui ne s'est guère conservée qu'en osmanli.

Mentionnons ici pour mémoire le *Codex comanicus*, vocabulaire latin-persan-turc (romanisé), datant du commencement du XIVe siècle de notre ère, et conservé à Venise (Bibl. de Saint-Marc). Cet ouvrage, très important pour l'histoire de la langue turque, a été l'objet des travaux de Klaproth, Geza Kuun, Radloff, Bang et d'autres (voir Pelliot, *A Propos des Comans*, J. as., avril-juin 1920, p. 126-127) ; il aurait besoin d'être réédité.

Structure phonétique.

Aux particularités signalées plus haut (p. 188) à propos de l'ensemble des langues turques, mongoles et tongouzes, nous ajouterons ici quelques détails sur le phénomène d'*harmonie vocalique* qui reçoit dans les dialectes turcs son maximum de rendement. Il y a là un mécanisme qui tient du merveilleux par sa complexité, sa régularité et sa symétrie.

Les lois de l'harmonie vocalique portent toutes sur les mêmes huit voyelles : *ä*, *i*, *ö*, *ü* ; *a*, *ị*, *o*, *u*. Elles déterminent *dans le cas le plus favorable* les phénomènes suivants :

1° Assimilation entre voyelles antérieures (*ä*, *i*, *ö*, *ü*) d'une part et voyelles postérieures (*a*, *ị*, *o*, *u*) d'autre part.

2° Assimilation entre voyelles larges (*a*, *o*, *ä*, *ö*) et voyelles étroites (*ị*, *u*, *i*, *ü*).

3° Assimilation entre voyelles arrondies (*o*, *ö*, *u*, *ü*) et voyelles neutres (*a*, *ä*, *ị*, *i*).

Le premier de ces phénomènes agit sur l'ensemble du mot fléchi ou non fléchi et se traduit par cette règle qu'un seul et même mot ne peut contenir que des voyelles antérieures ou des voyelles postérieures.

Ex. : (osm.) *öl-dür-mä-yälim* « ne tuons pas »,
(osm.) *otur-ma-yal̦ịm* « ne restons pas assis ».

Les deux autres agissent seulement sur les voyelles des suffixes grammaticaux considérées dans leur rapport avec la *dernière* voyelle du mot à suffixer (rapport qui est régi également par l'assimilation d'ordre plus général signalée en premier lieu).

Il résulte du triple jeu de ces assimilations que la voyelle d'un même suffixe peut recevoir quatre timbres différents : *a*, *o*, *ä*, *ö*, si elle est large ; *ị*, *u*, *i*, *ü*, si elle est étroite.

Voici le schéma de ces variations :

DERNIÈRE VOYELLE DU MOT SUFFIXÉ	VOYELLE SUFFIXALE POSSIBLE	
	dans les suffixes à voyelle large	dans les suffixes à voyelle étroite
a ou *ị*	*a*	*ị*
ä ou *i*	*ä*	*i*
o ou *u*	*o*	*u*
ö ou *ü*	*ö*	*ü*

Exemples pour le datif (suffixe à voyelle large) et l'accusatif (suffixe à voyelle étroite) :

Osm. *qotšar* (turc or. *qotšgar*, *qotšqar* ; turc ancien *qotšñar*) « bélier », datif *qotšar-a* (anc. *qotšar-ga*) ; accusatif *qotšar-ị* (anc. *qotšar-ịg*) ;

osm. *ämäk* (anc. *ämg'äk*) « peine, travail », datif *ämg'äy-ä* (anc. *ämg'äk-g'ä*), accusatif *ämäy-i* (anc. *ämg'äg-ig*), etc.

Suffixes désinentiels du turc.

Pour la définition des suffixes désinentiels, voir plus haut, p. 191).

I. *Suffixes désinentiels des noms turcs.*

Les *substantifs* peuvent recevoir les suffixes désinentiels suivants (nous les énumérons dans l'ordre où ils se placent en cas de cumul) :

1	2	3	4
suff. du pluriel	suf. possessif	suf. casuel	suf. *-ki*

1 2 3 4

Ex. : *är-lär-imiz-dä-ki* « qui est (sont) chez (dans) nos hommes ».

1° Suffixe du pluriel. Le seul suffixe du pluriel actuellement vivant en turc est *-lär* (formes secondaires : *-tär*, *-där* et *-När*). Il se joint non seulement aux substantifs et aux pronoms (sauf, en principe, ceux de la 1[re] et 2[e] personne), mais aussi à la 3[e] personne des formes verbales.

Ex. : *yịł-łar* « les années » ;

k'äl-di-lär (osm. *g'äl-di-lär*) « ils sont venus ».

2° Suffixes possessifs. Les suffixes possessifs, d'origine pronominale, figurent dans le tableau de la p. 210. Ils ont pour équivalents français les adjectifs possessifs « mon, ma ; ton, ta ; son, sa ; notre ; votre ; leur ».

Ex. : *qịz-ịm, qịz-ịṅ*, etc., « ma fille, ta fille, etc. ».

3° Suffixes casuels.

Les langues turques ont développé une déclinaison à petit nombre de cas. Ce sont :

génitif : *-(n)iṅ* [1], *-niṅ* ; formes secondaires : *-diṅ*, *-tiṅ* ;

accusatif : *-ig*, *-äg*, *-(y)i*, *-ni* ; formes secondaires : *-di*, *-ti* ; yak. aussi : *-nä*, *-tä*, *-lä* ;

datif : *-gä*, *-kä*, *-(y)ä* ;

locatif : *-dä*, *-tä* ; yak. *-nä*, *-inä* ;

ablatif : *-dän*, *-din* ; les mêmes avec *ṅ* au lieu de *n* ; les mêmes avec *t* au lieu de *d* ; *-nän*, *-näṅ*.

On peut ajouter à cette liste l'ancien cas instrumental en *-in* et le comparatif ou relatif dont les formes variables présentent l'intérêt de nous montrer comment des mots autonomes tels que

1. Il est possible qu'à l'origine le génitif et l'accusatif n'aient formé qu'un seul cas et qu'ils se soient différenciés postérieurement.

tän (*dän*, *tay*) et *tšag*, exprimant tous deux une idée de « quantité », sont devenus des suffixes casuels (*-täg*, *-däg*, *-täk*, *-däk*, *-täy*, *-däy*, *-tiy*, *-diy*; *tšäk*, *-džäk*, *tšä*, *-džä*, *-tšik*, *-džik* ; voir J. Deny, *Grammaire de la langue turque*, § 930 et 918).

Anciennement le locatif et l'ablatif ne formaient qu'un cas (suffixe *-dän*, probablement). Dans le turc de l'Orkhon, ils ne sont pas encore tout à fait différenciés.

La marque caractéristique des désinences casuelles par rapport aux autres particules consiste dans le fait qu'elles développent un *n* « pronominal », lorsqu'elles sont affectées à un mot déjà muni du suffixe possessif de la 3[e] personne : *itš* « intérieur », *itš-i* « son intérieur » ; à l'accusatif : *itš-i-n-i*.

On sait (voir p. 190) que le substantif peut figurer dans la phrase turque sans aucune désinence casuelle.

Il en est ainsi (en osmanli par ex.) lorsque le substantif (par ex. : *tavuq* « poule ») est :

a) Sujet de la proposition.

Ex. : *atš tavuq düš-ün-dä darị g'ör-ür* « poule affamée voit (du) millet dans son rêve » (proverbe).

b) Complément direct indéterminé.

Ex. : *bir tavuq sat-maq* « vendre une poule » ; *tavuq sat-maq* « vendre des poules, être marchand de poules, m. à m. poule vendre ».

Voir aussi le mot *darị* « millet » dans le proverbe ci-dessus.

c) Complément nominal, dans les expressions fixées par l'usage.

Ex. : *tavuq g'öks-ü* « sorte d'entremets, proprement : poitrine de poule » (le mot complété est muni d'un suffixe possessif servant d'indice de rappel au complément : m. à m. « poule sa poitrine »).

Le complément direct déterminé (voir plus haut, *b*) s'exprime par l'accusatif.

Ex. : *tavug-u* (pron. *tavu'-u*) *g'ätir* « apporte la poule ».

Le complément indirect prend la forme du datif ou de l'ablatif. Ces deux cas s'opposent l'un à l'autre : on met au datif les noms désignant les objets *vers* lesquels tend une action (= alla-

tif), et à l'ablatif ceux dont s'éloigne ou émane une action (au propre ou au figuré) : *su-ya baq-maq* « regarder l'eau (dat.) »; *su-dan qorq-maq* « craindre l'eau (abl.) ».

Ces mêmes deux cas joints au locatif jouent un rôle important comme cas locaux ou adverbiaux (compléments circonstanciels de lieu et de temps).

Le complément nominal (voir plus haut, *c*) dans les expressions non fixées par l'usage se met au génitif. Le mot complété garde son suffixe possessif.

tavug-uñ (*tavu'-uñ*) *g'öks-ü* « la poitrine de la poule » m. à m. « de la poule sa poitrine ».

Cette construction qu'on peut appeler « état construit » ou « rapport d'annexion » offre la particularité d'exprimer deux fois la possession (par le génitif et par le suffixe possessif).

4° Suffixe *-ki, -gi*.

Il ne se joint qu'au locatif ou au génitif des substantifs et à certains adverbes de lieu ou de temps, avec le sens de : « (celui) qui est dans, (celui) qui est de (à)... ».

Ex. : *yazi-da-ki* (*-qi*) « (celui, celle) qui est dans la plaine ».

Ce suffixe offre la particularité d'échapper fréquemment à l'application de la loi d'harmonie vocalique (il conserve par conséquent un aspect fixe). Bien qu'étant, à ce qu'il semble, d'origine verbale (voir *Grammaire turque,* § 1361), le suffixe *-ki* a été assimilé par la langue à un élément pronominal et développe un *n* à la flexion, comme les suffixes possessifs de la 3ᵉ personne.

En outre, il peut être appelé, de même que ces derniers, suffixe mixte (voir p. 191).

II. — *Copule turque.*

Les propositions nominales du turc se passaient à l'origine, et se passent encore souvent de la copule. Le schéma de ces propositions est donc : sujet + prédicat, c'est-à-dire l'inverse du schéma qu'il faudrait adopter pour les mêmes mots, si l'on voulait en faire une épithète + nom déterminé (voir p. 192).

Ex. : *ogłan ögsüz* (*öksüz*) « le garçon est orphelin » ;
ögsüz ogłan « un (ou le) garçon orphelin ».

De même avec les pronoms personnels, on dira :
bän ögsüz, sän ögsüz, biz ögsüz, siz ögsüz « je suis orphelin(e), tu es orphelin(e), nous sommes orphelin(e)s, vous êtes orphelin(e)s ».

La copule a paru assez tardivement ; ce n'est pas au présent du verbe « être » qu'on en emprunta les formes : on se servit de l'aoriste du verbe *dur-maq* (*tur-maq*) « rester debout, se tenir debout » (comparer latin *stare*). Cette construction fut affectée surtout à la 3ᵉ personne :
ogłan ögsüz tur-ur « le garçon est orphelin ».

Pour les 1ʳᵉ et 2ᵉ personnes, on se contenta de répéter, en le plaçant après le prédicat, le pronom personnel, mais cette fois comme enclitique.
bän ögsüz-bän, sän ögsüz-sän, etc.

Le *Qutadgu Bilig* et le *Divan* de Kachghari nous apprennent qu'il en fut de même anciennement à la 3ᵉ pers. : *oł ögsüz oł*.

En sa qualité d'élément inaccentué (l'accent tonique porte sur le nom du prédicat), l'enclitique obéit aux lois de l'harmonie vocalique comme tous les suffixes et, dans certains dialectes, se modifie en s'abrégeant.
bän ögsüz-üm (pour les autres formes, voir le tableau p. 210).

On peut considérer le pronom personnel enclitique comme un suffixe accompagnant le prédicat et l'appeler *suffixe prédicatif*.

L'emploi du suffixe prédicatif rend inutile, ou tout au moins facultatif, le pronom personnel sujet (autonome), à moins qu'il ne s'agisse d'exprimer l'insistance : « c'est moi qui suis orphelin » : (*bän*) *ögsüz-üm*.

D'autre part, la langue use parfois aussi de la faculté de rendre à la copule un peu de son ancienne indépendance en intercalant entre elle et le prédicat certaines particules (en osmanli par ex. l'interrogatif *mi*, le comparatif *gibi* et le négatif *deyil*) :

ögsüz-mü-bän }
osm. — *-mü-yüm* } « suis-je orphelin ? »

TABLEAU DES DÉSINENCES D'ORIGINE PRONOMINALE

(Y COMPRIS, A LA 3e PERSONNE, QUELQUES DÉSINENCES D'ORIGINE VERBALE).

	Pronoms personnels autonomes.	Copule ou désinences prédicatives = pronoms personnels enclitiques.	Désinences de l'impératif.	Désinences prédicatives spéciales.	Suffixes possessifs.
			SINGULIER		
1re pers.	bän, män, min	-bän, -vän, -män, -än, -äm -bin, -min, -in, -(y)im	-äy, -äyin, -(y)äyim -iyin, -iyim, īm	-m	-[i]m
2e pers.	sän, sin	-sän -sin ; yak. -gin et -kin	Zéro (ou -kil, -kin -kilän, -kinän)	-ñ, -g	-[i]ñ
3e pers.	ol, ul, o ; fléchi : an-, on-	-ol, -ul (en turc anc.) -durur, -dir, -di ; les mêmes avec t, au lieu de d ; -t	-sin, -zin, -zü ; yak. : -tin, -din, -lin, -nin	Zéro	-(s)i, -(z)i ; yak. -(t)ä
			PLURIEL		
1re pers.	biz, bis, pis	-biz, -bis, -pis ; yak. -bit ; -miz, -(y)iz, -(y)ik	(y)älim -äylik, -älik, -äli, -äl, -äligär, etc. -äni, -äñ -(y)äyik, -iyik, -äk, yak. -iäx	-bis, -miz ; yak. -bit -k	-[i]biz, -[i]miz ; yak. -[i]bit
2e pers.	siz, siz-lär, silär	-siz, -siz-lär, -silär, -siñnär, -siñär, -sär -siñiz, yak. -git	-(y)iñiz, -igiz, -[i]ñ, -(y)iñ -[i]ñär, -[i]gär	-ñiz, -giz, yak. -git -ñär, -gär	-[i]ñiz, yak. -[i]gi -[i]ñär, -[i]gär
3e pers.	olar, ular, onlar, onnar, anlär	Les mêmes suffixes que pour le singulier ou les mêmes suivis de -lär (sauf pour le suff. poss. : -läri à côté de -(s)i.			

III. *Suffixes désinentiels du verbe turc.*

Ils sont de deux sortes :

1° Suffixes des formes nominales du verbe. Ce sont des morphèmes de nature mixte (voir p. 191). Tels sont par exemple, parmi les plus importants, le suffixe *-kän* et son synonyme osmanli *-dik*.

Ex. : *al-gan* (osm. *al-diq*) « fait de prendre [nom d'action], (celui) qui prend [participe], (celui) que (quelqu'un) a pris [participe quasi passif. »]

Les formes nominales du verbe turc servent dans une mesure particulièrement considérable à faire jouer au verbe le même rôle dans la phrase que le nom lui-même, c'est-à-dire à l'employer, suivant les formes, comme une sorte de substantif (= noms d'action), d'adjectif (= participes ou pro-participes), ou d'adverbe (= gérondifs et locutions gérondives).

Cette différenciation ne s'est opérée d'ailleurs que progressivement et pas pour toutes les formes, et l'on trouve d'abondantes traces (notamment pour les suffixes *-kän* et *-dik*) d'un état plus ancien où *la même forme nominale* servait tantôt de substantif, tantôt d'adjectif ou même d'adverbe.

Lorsqu'il s'agit d'un nom d'action, le sujet logique de la forme nominale du verbe prend la place d'un complément nominal de cette forme (le sujet possède donc l'action plutôt qu'il ne l'exerce, dans cette conception). Au lieu de dire, comme le français : « je sais que le meunier dort », le turc dit « je sais le fait (l'action) de dormir du meunier » ou plus exactement : « je sais du meunier (au génitif) son fait de dormir (à l'accusatif) » : osm. *däyirmändži-niṅ* (anc. *tägirmän-dži-niṅ*) *uyu-dug-u-nu* (anc. *udu-gan-i-ni*) *bil-ir-im*.

Le groupe de mots « du meunier son fait de dormir » comporte implicitement un jugement — comme les propositions — à savoir : « le meunier dort », mais comme construction et comme place dans la proposition, il ressemble à un terme ordinaire de celle-ci et porte la marque de l'accusatif comme n'importe quel substantif servant de complément direct, par ex. : *diley-iṅ-i bilirim* « je connais ton désir ».

On peut donc appeler de pareils groupes de mots « quasi-propositions ». Il peut y avoir des quasi-propositions substantives, adjectives et adverbiales.

2° Suffixes des *formes personnelles* du verbe ou *formes verbales proprement dites* (comme elles représentent toujours le prédicat d'une proposition verbale, on peut les appeler aussi formes *prédicatives*).

Ces suffixes sont les suivants : *a*) les désinences de l'impératif; *b*) les désinences des autres formes personnelles.

a) Les désinences personnelles de l'impératif diffèrent de celles des autres formes (voir le tableau p. 210). L'impératif est, en outre, la seule forme dont les désinences personnelles se joignent directement à la base verbale.

Ex. : *at-ịñ* ou *at-ịñịz* « jetez ».

b) Les suffixes désinentiels des autres formes personnelles sont complexes. Ils comportent : α) un signe thématique, β) une désinence personnelle (ou prédicative).

α) Les suffixes que nous appelons, à défaut d'autre terme commode, « thématiques », portent les notions de temps et de mode. Ce sont, en réalité, des suffixes de formes nominales (de participes), anciens ou encore vivants. En s'ajoutant à la base, ils forment un « thème ».

Ex. : *säv-miš* = 1° participe passé, « qui a aimé »; 2° thème du passé indéterminé.

Le nombre des thèmes varie suivant les dialectes. En osmanli, il y en a huit.

β) Les désinences personnelles (prédicatives) qui se joignent à un thème sont en général des *pronoms enclitiques*.

En effet, presque toutes les formes personnelles actuelles étaient anciennement des *propositions nominales* dont le prédicat (représenté par un participe) est devenu thème et dont la copule est devenue désinence personnelle.

On retrouvera donc ici le même processus que ci-dessus (p. 209) en remplaçant le mot *ögsüz* par le mot *säv-miš*, par exemple. On a dit successivement : **bän säv-miš*; *bän säv-miš-bän* (osm. *bän säv-miš-im*); *säv-miš-bän* (osm. *säv-miš-im*), toutes formes qui correspondent à nos « j'ai aimé, j'aimai ».

Il résulte de ce qui précède que les désinences personnelles peuvent être séparées du thème par les mêmes particules que la copule l'est de son prédicat (en osm. les particules *mi*, *gibi*, *deyil*) : *säv-miš-mi-im* « aimai-je ? ».

Il existe cependant deux thèmes verbaux — celui du prétérit en *-di* et celui du conditionnel en *-sä* — qui prennent comme désinence personnelle ou prédicative non la copule, mais une désinence spéciale qu'on peut appeler désinence prédicative spéciale ou proprement verbale.

Ex. : *säv-di-m*, *säv-di-ñ*, *säv-di...* « j'ai aimé, j'aimai ; tu as aimé, tu aimas ; il a aimé, il aima, etc. » ;

säv-sä-m, *säv-sä-ñ*, *säv-sä* « si j'aime (aimais), etc. ».

Ces deux formes, dont la première se trouve déjà employée dans les textes les plus anciens, paraissent néanmoins les plus évoluées, puisqu'elles sont les plus affranchies du caractère de forme « nominale » ancienne.

La désinence s'incorpore ici étroitement au thème dont aucune particule ne peut la séparer. Elle est également d'origine pronominale, mais ressemble davantage à des suffixes possessifs (voir le tableau p. 210), ce qui, à moins qu'il ne s'agisse d'une simple contamination, permet de supposer que les formes en *-dim*, *-diñ*, etc. sont des vestiges d'une conjugaison analogue à celle dont on constate l'existence en finno-ougrien et qu'on appelle conjugaison subjective (voir aussi C. Brockelmann, *Keleti Szemle*, 1918).

Indépendamment des formes personnelles examinées jusqu'ici et qu'on peut appeler formes simples, il existe des formes composées (= thème + formes du verbe substantif autres que le présent). Ce sont aussi d'anciennes propositions nominales. Les formes personnelles dont le nombre se trouve ainsi fortement augmenté peuvent, grâce à ce procédé, exprimer les notions modales ou temporelles et des notions d'aspect très complexes et très nuancées. L'osmanli possède trois formes composées principales (passée, suppositive et dubitative) pour chacun de ses thèmes.

Remarque. — Il résulte de l'exposé qui précède que la catégorie grammaticale du nom est très importante en turc et qu'elle domine

le verbe dont les formes personnelles elles-mêmes portent des traces de formation nominale (traces devenues presque imperceptibles dans les deux thèmes en *-di-* et *-sä-* dont il vient d'être parlé).

Il ne faudrait pas en conclure, cependant, que la notion verbale se confond ici avec la notion nominale. Le vocabulaire turc distingue nettement les racines verbales et les racines nominales. On n'y trouve qu'un nombre restreint de mots pouvant servir à la fois de nom et de verbe. Si le même groupe de phonèmes assume ce double rôle en turc, c'est qu'il s'agit, presque toujours, d'une simple rencontre. Ex. : *at* « cheval » ; *at* ! « jette ! »

Écriture.

Les Osmanlis et les Turcs, en général, se servent, en leur qualité de Musulmans, de l'alphabet coranique, c'est-à-dire de l'alphabet arabe, légèrement modifié par les Persans (signes diacritiques ajoutés aux lettres *bā*, *dẑim* et *ẑā* pour figurer les sons *p*, *tš* et *ž*).

Cet alphabet ne suffit pas à noter les nuances du turc. Ainsi pour les quatre voyelles labiales *o*, *ö*, *u*, *ü*, on ne dispose que d'un seul signe, le *vav*. Pour transcrire le nom de la rivière Vouvo, les géographes turcs alignent 4 *vav* !

Avant d'avoir adopté l'alphabet arabe, les Turcs ont connu deux écritures qu'on peut considérer comme nationales. Elles sont d'origine sémitique (aussi y retrouve-t-on l'imperfection du système de notation des voyelles). Ce sont :

1° L'alphabet RUNIFORME (ainsi appelé à cause de sa ressemblance extérieure avec les runes) ou *kök-türk* ou de l'Orkhon.

2° L'alphabet OUÏGOUR OU NÉO-SOGDIEN.

L'alphabet runiforme a été en usage dans l'Empire des T'ou-Kiue (voir p. 194) et chez les Kirghizs de l'Iénisseï. Il était déjà en usage au VIe s. de notre ère, et bien que les alphabets dont il a été dérivé directement ne soient pas encore connus, Donner le rattache à une écriture araméenne plus ancienne que l'alphabet des Arsacides, et qui devait être employée dans « les premiers

siècles de notre ère presque partout dans le Turkestan » (*Sur l'origine de l'alphabet turc*, page 44). Elle ressemble à celle qu'on trouve « sur les monnaies des satrapies de l'Asie-Mineure, de la Mésopotamie, de la Cilicie et plus tard dans les papyrus d'Égypte »

Les caractères s'écrivent sans ligature, isolément, la fin des mots ou groupes de mots étant marquée par deux points (:).

Cet alphabet qui se compose d'une quarantaine de lettres tient compte au plus haut degré du phénomène de l'harmonie vocalique : les sons *b*, *d*, *g*, *y*, *k*, *f*, *s*, *r*, *l*, *n* sont représentés par des lettres différentes suivant qu'ils figurent dans un mot à voyelles postérieures ou antérieures (dans l'alphabet arabe-turc on n'a des doublets que pour les gutturales ; en osmanli, pour les dentales également).

Quatre signes seulement figurent les voyelles dont la graphie est souvent omise non seulement à l'intérieur du mot mais parfois même à l'initiale.

Le runiforme s'écrit de droite à gauche.

Quant à l'alphabet ouïgour, la nature exacte en a été reconnue par Klaproth en 1811 [1]. Il dérive, ainsi que l'ont établi en 1909 F. K. W. Müller et, à sa suite, R. Gauthiot et A. von Le Coq, de l'alphabet sogdien et non, comme on le croyait avant, de l'estranghelo syrien des religieux nestoriens.

On ne connaît pas la date à laquelle cette écriture a été introduite chez les Turcs ; elle paraît contemporaine des textes runiformes les plus récents (A. von Le Coq, *Kurze Einführung in die uigurische Schriftkunde*, Mitt. Sem. Berlin, 1919, p. 93 à 109).

L'écriture ouïgoure a servi à la rédaction de textes bouddhiques et, avant l'introduction définitive de l'alphabet arabe, de textes musulmans.

Les découvertes faites par diverses missions au Turkestan chinois ont montré que les Turcs se sont servis également des écritures suivantes :

1. On a cependant retrouvé les traces d'un déchiffrement opéré dès le XVII[e] siècle (voir Pavet de Courteille, *Mirādj-Nameh*, p. XV).

le *manichéen* (ou estranghelo modifié) qui a été déchiffré par F. K. W. Müller était employé pour les textes religieux des Manichéens (la même écriture servait à écrire les textes manichéens en langue sogdienne).

le *syriaque* (estranghelo) employé pour les textes chrétiens-nestoriens (inscriptions tombales publiées par Chwolson).

l'écriture *brahmi* (voir à l'index).

l'écriture *tibétaine* (voir à l'index).

De nos jours les Arméniens de langue turque, les Grecs de Caramanie et les Caraïtes se servent respectivement des écritures arménienne (qui s'adapte bien au turc), grecque et hébraïque.

L'Azerbeïdjan du Caucase a adopté récemment les caractères latins.

Langue écrite et littérature.

Les monuments les plus anciens de la langue turque appartiennent à l'épigraphie. Ils sont, comme on l'a vu plus haut, du VIII[e] siècle de notre ère. Deux ou trois siècles plus tard, les Turcs possédaient une littérature religieuse, — surtout de traductions — bouddhique, manichéenne et chrétienne, dont de précieux spécimens, généralement de date incertaine, ont été mis au jour par les missions allemandes de Grünwedel (1902-1903), de von Le Coq (1904-1905) et de ces deux savants réunis (1905-1907), la mission anglaise d'Aurel Stein (1906-1908), la mission française de Paul Pelliot (1906-1909), et la mission japonaise de M. Tachibana (1908). A Kachghar le chambellan (khass-hadjib) Youssouf écrivit en 1069 le *Qutadgu Bilig*, ouvrage didactique en caractères ouïgours (publié par Vámbéry et par Radloff), et vers 1074, Mahmoud Kachghari composa en arabe le *Divānu-lugat-it-turk* ou « Recueil de mots turcs », dont il a été question plus haut, imprimé pendant la guerre à Constantinople et qui nous livre des spécimens (en caractères arabes) de la littérature populaire et didactique de cette époque.

Jusqu'au XIV[e] siècle les ouvrages littéraires turcs sont peu nombreux. Au XIII[e] siècle la littérature « seldjoukide » de Roum ou d'Anatolie est pauvrement représentée. La littérature otto-

mane apparaît dès la fondation de l'Empire (1300) avec Achiq Pacha mort en 1310-1311. Elle a fourni un très grand nombre d'écrivains. Depuis 1850 cette littérature s'est affranchie de l'influence persane et arabe, pour subir celle de la littérature européenne (surtout française : apparition du roman et du théâtre), et en ce moment elle cherche à dégager son génie national.

L'imprimerie existe depuis 1730.

En Asie Centrale régna une autre langue littéraire, contemporaine, en somme, de l'osmanli et plus ou moins artificielle, appelée tchaghataï (du nom du second fils de Gengiskhan qui eut pour apanage cette région). Elle fleurit à Samarcande, Boukhara, Andidjan et Hérat et fut illustrée aux XVe, XVIe et XVIIe siècles par Mir Ali chir Nevaï, le sultan Baber et Aboul-Ghazi Behadour Khan. On y trouve, rarement, il est vrai, des ouvrages en écriture ouïgoure (voir le *Tezkere-i- Evliya* traduit du persan et le *Miradj-Name*, publiés par Pavet de Courteille). Cette littérature exerce encore son influence — qui cède cependant de plus en plus le pas à celles des dialectes locaux et de l'osmanli de Constantinople — chez les peuples turcs qui se sont mis à écrire leur langue : à Kazan, dans les Républiques récentes de Boukhara, Khiva, Turkestan, Azerbeïdjan et Crimée. Les Kirghizs eux-mêmes suivent le mouvement.

Force d'expansion.

Nous ne connaissons que peu d'exemples de peuples turcs ayant abandonné leur langue (le plus important est celui des Bulgares). On a, par contre, vu plus haut (p. 187) plusieurs exemples de peuples turquisés, auxquels on peut ajouter celui des Sartes (proprement des « marchands ») qui du temps de Névaï parlaient encore persan. Les khans mongols de la Horde d'or qui domina en Russie n'usaient que du turc comme langue administrative et Timourleng parlait turc. Le koumik sert de langage international aux populations voisines du Daghestan et ne perd ce rôle que devant l'influence grandissante du russe.

En Turquie même, où, contrairement à ce qui se passe sou-

vent ailleurs, les questions de langue ne se confondent pas avec les questions de religion, les catéchismes orthodoxes des Grecs de Caramanie sont rédigés en turc (en caractères grecs). Beaucoup d'Arméniens sont non seulement bilingues, mais préfèrent — les hommes surtout — parler le turc, et c'est dans cette langue que leurs prêtres lisent fréquemment l'évangile à la messe. Une partie des Arméniens de Pologne, aujourd'hui complètement polonisés, ont parlé turc (du XIVe au XVIIe siècle) (voir *Journal Asiatique*, juillet-sept. 1921, p. 135).

Dans les Balkans où les Osmanlis, malgré l'allure superficielle en quelque sorte et extérieure de leur domination, ont hérité de l'hégémonie de Byzance, une foule de mots ont été introduits ou véhiculés par eux. Ce sont, pour la plupart, des formules de politesse, des termes d'étiquette et des substantifs relatifs à la vie quotidienne, vêtements, cuisine, mobilier, produits d'industrie et autres objets dont l'ensemble donne l'impression d'une civilisation « balkanique » et « orientale » allant de la Morée jusqu'en Bosnie et Roumanie, survivance plus ou moins modifiée et turquisée de la civilisation byzantine. Mochkof a compté 8,7 p. 100 de mots turcs dans le dictionnaire bulgare de Duvernoy. Il en aurait trouvé bien plus en dépouillant le vocabulaire des chants nationaux bulgares ou serbes. Ce n'est que depuis qu'elles ont reçu leur indépendance que les populations balkaniques éliminent, souvent avec succès, cet apport étranger.

BIBLIOGRAPHIE

On trouvera une liste des principaux ouvrages concernant le turc en tête de la Grammaire turque que nous avons publiée en 1920 chez Ernest Leroux.

Les travaux de M. V. Thomsen sont maintenant réunis dans le volume III de ses *Samlede afhandliger* (Copenhague, 1922).

Sur le tchouvache, voir, en dernier lieu : Ramstedt, *Zur Frage nach der Stellung des Tschuwassisches*, Helsingfors, 1922 (*Journal de la Société finno-ougrienne*, XXXVIII).

LANGUES MONGOLES

Les Mongols.

Les origines des Mongols qui s'appellent eux-mêmes *Mongol* ou *Mogol* sont au moins aussi obscures que celles des Turcs (voir plus haut, p. 186). Leur nom apparaît pour la première fois dans les annales chinoises de la dynastie des T'ang (618-907) qui mentionnent le peuple des Che-wei Mong-Wou (ou Mong-Wa). Dans les annales de la dynastie des Song x^e au xiiie s.), on les retrouve sous le nom de Mong-Kou (Mong-Kou-sseu) et au xiie siècle ils forment un état qui fait la guerre aux Kin (ou Niu-Tchen, voir p. 234) en 1135, 1147 et 1161. On suppose que les Mong-Kou avaient réussi à s'imposer aux autres tribus du peuple Che-wei qui nomadisait dans le Nord et l'Est de la Mongolie actuelle.

Avant ces Mong-Kou ou à la même époque qu'eux, ont surgi dans l'histoire d'autres peuples que soit la généralité des savants, soit certains d'entre eux rattachent aux Mongols. C'est ainsi que sans reparler des énigmatiques Hiong-Nou (voir p. 186), on peut mentionner les Sien-pi qui, au milieu du iiie siècle de notre ère, émigrèrent de la région du fleuve Leao (ou Liaou-ho en Mandchourie) pour s'installer dans le Nord du Kan-Sou. Ils sont des Tongouzes pour Chavannes, Parker et Franke et des Mongols pour Pelliot, Rockhill et Laufer.

Aux Sien-pi se rattachent les T'ou-yu-houen qui allèrent fonder un royaume dans la région du Koukou-nor au début du ive siècle.

Les Jouan-jouan ou Avars, vaincus par les T'ou-Kiue (voir

p. 194) et qu'il ne faut pas confondre avec les Pseudavares ou Ouarkhonites qui, battus en même temps, allèrent fonder un empire dans la Hongrie actuelle, sont considérés, généralement, comme ayant été de langue mongole.

Les K'i-tan [1] qui furent les maîtres de la Chine septentrionale (dynastie Leao : 907-1125) et auxquels on attribuait jusqu'ici une origine tongouze, parlaient, en réalité, d'après M. Pelliot, « une langue étroitement apparentée au mongol, encore que fortement palatalisée » (*A propos des Comans*, p. 22).

Mais ce n'est qu'au commencement du XIIIe siècle que les Mongols proprement dits font leur entrée, de plain-pied, dans l'histoire, d'une manière foudroyante d'ailleurs. Leur chef Temoutchine se fait proclamer en 1206 (?) empereur sous le nom de *Tšiṅgiz-Xan* (= Gengiskhan). Sous ce prince et ses premiers successeurs, les bannières mongoles pénétrèrent un moment jusqu'à l'Oder, l'Adriatique, jusqu'en Anatolie, en Syrie, dans le Tibet, en Indo-Chine, à Java, au Japon.

Cet immense empire, qui s'étendait de la Corée aux frontières de la Pologne, ne tarda pas à se désagréger.

En Chine, Khoubilaï-khan, l'héritier du titre de grand khan et aussi le premier empereur bouddhiste des Mongols, donna à sa dynastie le nom chinois de « Yuan » après avoir transporté la capitale mongole de Karakoroum à Péking (en turc *Xanbaliq*, la « ville khanienne »); mais dès 1368 ses successeurs en étaient chassés par la dynastie nationale chinoise des Ming.

Les autres dynasties fondées par les descendants de Gengiskhan et qu'un lien de vassalité de plus en plus lâche rattachait aux Yuan disparurent également dans le courant du même XIVe siècle. Tel fut le sort de la dynastie de Batou (fils de Djotchi, fils de Gengiskhan) dans le Kiptchak et la Russie, de la dynastie de l'il-khan Houlagou (fils de Toulouï, fils de Gengiskhan) en Perse, de celle de Tchaghataï (2e fils de Gengiskhan) dans le Turkestan.

1. De là l'arabe *ḫaṭā(y)* « Chine du Nord » et le russe *kitay* « Chine ». Comparer le Cathay des auteurs occidentaux.

L'empire, également éphémère, que reconstitua à la fin du XIVe siècle Timour ou Tamerlan, n'avait de mongol que les lointaines origines de son fondateur, devenu tout à fait turc et très musulman. Il en fut de même pour les petites dynasties qui se maintinrent encore assez longtemps en Kachgharie, en Crimée et dans l'Inde (les grands Mogols).

En Mongolie même, la puissance des Mongols finit par sombrer dans les discordes intestines. Les Mandchoux en profitèrent dans le cours des XVIIe et XVIIIe siècles pour se soumettre le pays auquel ils ont laissé cependant son organisation propre fondée sur une division — à la fois tribale, militaire et administrative — en *tsuglān* (diètes ou ligues), *aymaq* (tribus ou corps d'armée ; comparer le turc *oymaq*), *xošūn* (bannières ; comparer le turc *qošun*) et *sumun* (escadrons ou compagnies ; proprement « flèches »).

Cette autonomie relative semble appelée à bénéficier des événements actuels : les révolutions chinoise et russe tournent à l'avantage de ce peuple qui, sous l'influence du bouddhisme, pense-t-on, a entièrement perdu son caractère guerrier.

Distribution géographique de la langue mongole.

Contrairement à ce qui s'est passé pour les Turcs, l'immense extension qu'a atteinte à un moment donné l'Empire des Mongols n'a pas causé leur dispersion. A part les Kalmouks que des migrations relativement récentes (voir p. 187) ont portés sur les territoires appartenant à la Russie, à part aussi les Aïmaks de l'Afghanistan, la population mongole est pour ainsi dire centralisée en Mongolie, dont les quatre cinquièmes du territoire lui appartiennent, et dans les pays contigus.

Les Mongols sont d'ailleurs au moins dix fois moins nombreux que les Turcs. Leur nombre exact n'est pas connu, faute de statistique précise. Les chiffres adoptés par Deniker donnent 1.950.000 individus. Ils doivent être inférieurs à la réalité puisque pour les Bouriates seuls le recensement russe de 1897 en accusait 289.001 au lieu des 250.000 de Deniker. Rockhill parle de 1.800.000 Mongols pour la Mongolie seulement, ce qui, avec les

486.198 qui habitent la Russie, donnerait environ 2.280.000 individus. Ivanovski admet le chiffre global de 3.000.000 qui nous paraît se rapprocher davantage de la vérité. Il faudrait même ajouter à ce nombre les 5 ou 600.000 Mongols de l'Afghanistan, mais on ignore dans quelle mesure ils ont conservé leur langue.

D'après le même auteur la population de la Mongolie augmente, mais dans des proportions insignifiantes. Les Mongols vivent dans des conditions matérielles (misère), sanitaires (maladies vénériennes) et sociales (grand nombre de prêtres voués au célibat) défavorables au développement de la race. Quant aux Bouriates (Russie), le taux annuel de leur accroissement atteindrait d'après Patkanov 0.4 p. 100 en Transbaïkalie et serait moindre ailleurs.

Voici la classification des dialectes mongols (ou plutôt des tribus mongoles), principalement d'après Roudnev et Soulié (XIV[e] congrès des orientalistes d'Alger).

I. Dialectes orientaux ou méridionaux, les plus importants par le nombre des sujets parlants : Mongolie du Sud et de l'Est, confins chinois, c'est-à-dire territoire des *Tšaxar*, et différentes diètes de la Mongolie dite « Intérieure » (carte : 1 à 10) :

Tšaxar (carte : 1); [Diète de *Dẓerim* :] *Xortšin Dẓalait*, *Durbut-böysö* (carte : 2), *Gorlos* ; [Diète de *Dẓosotu* :] *Xaratšin* (carte : 3), *Tumut* ou *Tumed* (carte : 4); [Diète de *Tšu-uda* :] *Aoxan* (carte : 5), *Nayman* (carte : 6), *Dẓarot*, *Aru-xortšin*, *On(g)nüt* (carte : 7), *Kešixtön*, *Barin* (carte : 8) ; [Diète de *Silingol* :] *Udẓumtsin*, *Xutsit*, *Sunit* (carte : 9), *Abagas* ; [Diète de *Ulan-tšab* :] *Durbön-xuxöt*, *Mu-mingan*, *Urot* ; [Diète de *Yökö-tsu*] : Ordos (carte : 10) ; *Suruq*.

II. Dialectes *xalxa* ou *xalxas*. Ils occupent un vaste territoire dans la Mongolie du Nord, moins le Nord-Ouest alias « Mongolie Extérieure » (carte : 11 à 14) : *xalxa* oriental (aïmak de *Tsetsön-xan* ; carte : 11); *x*. d'Ourga (aïmak de *Tušetu-xan* ; carte : 12); *x*. occidental (aïmak de *Dẓasagtu-xan* ; carte : 14), de *Sain-Noyon* (aïmak du même nom ; carte: 13) ; *x*. du Sud-Est (il existe une bannière de *xalxa* dans la diète le *Tšu-uda*, une

autre dans celle d'*Ulan-tšab*). On trouve aussi des *xalxa* en Dzoungarie et dans la région de Koukou-nor.

III. Dialectes bouriates (carte : 15 à 17). Principalement en Sibérie où les Bouriates étaient au nombre de 289.001 en 1897 et 332.554 en 1911 :

1. bouriate cisbaïkalien ou b. d'Irkoutsk (dialectes de Tounkinsk, Alarsk, Koudinsk, Balagansk, Kapsalsk, Nijné-Oudinsk, Idinsk) (carte : 15).

2. bouriate transbaïkalien (dialectes de Selenguinsk, Bargourinsk, Tsongol, *Xori*, Aguinsk, Koudarine) (carte : 16).

3. bouriate-Bargou (*Barxu*) en Mandchourie, (carte : 17).

IV. Dialectes kalmouks (Région d'Astrakhan, Caucase et Mongolie occidentale où les Mongols sont appelés Dzoungars) (carte : 18).

Le mot Kalmouk est un terme russe tiré de *xalimag*, surnom des Mongols installés en Russie d'Europe. Les Kalmouks de Russie sont au nombre de 190.648 dont 14.812 au Caucase (provinces de Stavropol et de Terek surtout), et 1.825 (en 1911 au lieu de 2.849 en 1897), en Asie Centrale. En Russie d'Europe les Kalmouks ont été décimés par les récentes famines.

En Dzoungarie, on trouve les *Dörböd* au S. E. des lacs d'*Atšit-nor* et d'Oubsa, les *Xoit* (au S. E. du lac d'Oubsa et au confluent du Sakosaï et du Kobdo), les *Xošod* (E. de la riv. Tchinguil), les *Tšaratšin* (E. de la riv. Boulgoun), les *Mingit* (Nord de la ville de Kobdo) ; les *Torgud* (carte : 18 *a*), sur l'Ouroungou ; les *Ölöd* (*Olñt*, Eleutes ; carte : 18 *b*) à l'Ouest de la ville de Kobdo ; les *Bait*. On trouve aussi des *Xoit* (4 bannières) et des *Xošod* (21 bannières) dans la région du Koukou-nor, comme on trouve des *Torgud* et des Eleutes, dans l'Alachan (à l'Ouest de la boucle du Fleuve Jaune). — On appelle parfois abusivement Eleutes les Mongols occidentaux en général, et on confond ce mot avec *Oyrad* « confédérés », nom qui désignait au XVIIIe siècle, par exemple, un groupement de 4 tribus (*Dörbön oyrad*) : les *Xošod*, les *Dörböd*, les *Dzungar* et les *Torgud*.

V. Mongols du Nord-Est du Tibet (carte : 19).

Il existe 8 bannières de Mongols dans la région de Tsaïdam.

En outre, dans la région du lac de Konkou-nor jusqu'aux rives du Fleuve Jaune, on trouve les *Tšoros*, et les *Xošod* et *Xoit*, déjà nommés.

VI. Mongols de l'Afghanistan (carte : 20), au nombre de 518.000 environ, en 1910 (d'après l'État-Major Gén. des Indes). Appelés Hezara, Aïmak, etc., mais se donnant à eux-mêmes le nom de Mongol, ils occupent une étroite bande de terrain qui va des environs de Hérat à la frontière indienne (qu'elle dépasse légèrement en passant par Kaboul. Sont en train d'abandonner leur langue pour le persan et le pouchtou (voir Ramstedt, *J. de la Soc. Finno-Ougr.* de 1905). Religion : musulmans chiïtes.

Indépendamment de ces dialectes, les Mongols possèdent une langue écrite ou littéraire dont on ne sait ni à quelle époque, ni dans quelles conditions elle s'est constituée. Aussi faut-il considérer la prononciation qu'on lui donne comme plus ou moins conventionnelle.

Les différents dialectes mongols se ressemblent beaucoup entre eux. Ils sont encore moins aberrants que les dialectes turcs.

L'une des différences les plus importantes consiste dans le traitement des sons *dž* et *tš* (de la langue littéraire).

Dans les dialectes orientaux (méridionaux) ces sons conservent leur prononciation en toute position, dans les dialectes khalkha (ou du Nord) seulement devant *i*. Devant les autres voyelles *dž* y est remplacé par *dz* et *tš* par *ts* :

Ex : khalkha *tsagān* × mong. oriental *tšagān* « blanc ».

En kalmouk *dz* devient *z*.

En bouriate les affriquées *tš* et *ts*, à l'initiale, se réduisent respectivement en *š* et *s*.

De même qu'en turc, la dentale explosive est généralement sourde à l'initiale (*t*), dans les dialectes orientaux, et sonore (*d*) dans les dialectes occidentaux (kalmouk, khalkha occidental). Elle est sonore en bouriate.

Aux explosives *k* et *q* de la langue littéraire correspond la continue *x* (dialectes orientaux), sauf à la fin des mots.

Structure phonétique.

Aux quelques indications qui précèdent on peut ajouter encore les suivantes.

La grande loi de l'harmonie vocalique reçoit, comme en turc, de fréquentes entorses. En principe, les mots de la classe postérieure ne doivent contenir que les voyelles postérieures *a*, *o*, *u* et parmi les consonnes gutturales : *q* et *g*. Par contre les mots de la classe antérieure ne doivent contenir que les voyelles (antérieures) *ä*, *ö*, *ü* et parmi les consonnes gutturales *k* et *g'*. Contrairement à ce qui se passe en turc, la voyelle *i* est indifférente, c'est-à-dire qu'elle se retrouve aussi bien dans les mots de la classe postérieure que dans ceux de la classe antérieure. Certains indices permettent de supposer qu'anciennement il existait une voyelle postérieure *į* pour les mots de la classe postérieure.

Les voyelles longues existent dans la langue parlée, mais la langue littéraire les ignore et offre toujours à leur place deux voyelles séparées par une gutturale (quelquefois par la labiale *b*) : *aga* = *ā* ; *ogo* = *ō* ; *ugu* = *ū* ; *agu* = *ū* ; *ägü* = *ǖ*, etc.

Bien que, dans certains cas, la langue littéraire représente l'ancien état de choses (voir p. 203 un phénomène analogue dans les dialectes turcs du Sud, par réduction de la gutturale intervocalique : *aga* > *aga* > *a'a* > *ā*), les mongolisants estiment que dans la plupart des cas, la graphie littéraire est ici purement conventionnelle. Parfois aussi la gutturale littéraire remplace abusivement une ancienne labiale disparue.

Le *b* intervocalique se prononce presque sans joindre les lèvres, presque comme *w*.

Suffixes désinentiels du mongol [1].

A la différence du turc, le mongol ne procède pas toujours par simple « agglutination » de suffixes. Quelquefois, — assez rarement, il est vrai, — il y a substitution ou remplacement d'un élément phonétique donné par un autre (alternance).

1. Voir page 191.

Ex. : *balgasun* « ville » (sg.) × *balgat* « villes » (pl.)
nöx'ür « compagnon » × *nöx'üt* « compagnons ».
(d'où le persan *nügär* « domestique »).

Il en résulte qu'en mongol certains substantifs peuvent n'avoir jamais de forme « nue » (équivalente à la racine). En turc, *tout* mot peut se présenter sous cette forme.

En outre, le même suffixe désinentiel mongol peut avoir plusieurs formes dans le même dialecte (voir plus bas à propos du pluriel). Pareil phénomène est tout à fait exceptionnel en turc.

I. *Suffixes désinentiels des noms mongols.*

Les substantifs peuvent recevoir les suffixes désinentiels suivants (énumérés dans leur ordre en cas de cumul) :

1	2	3	4
suff. du pluriel	suff. casuel	suff. possessif	suff. *-gi* (ou *-x'i*).

Ces éléments correspondent, à peu de chose près, à ceux qu'on trouve en turc, mais l'ordre n'en est pas le même, le suffixe casuel se plaçant ici avant le suffixe possessif. Cette différence doit correspondre à une différence chronologique dans le développement des suffixes : le turc a sans doute commencé par les suffixes possessifs et le mongol par les suffixes casuels. Étant donnée la nature des suffixes possessifs en mongol (voir plus loin, p. 228) cette hypothèse n'a rien d'invraisemblable.

1° Suffixes du pluriel. Il y a plusieurs suffixes du pluriel en mongol : *-nār*, *-nü'üt* (moderne *-ūt*), *-t*, *-s*. L'emploi de ces différentes désinences se détermine par la nature de la finale du nom, par le nombre des syllabes, le sens (êtres animés ou inanimés) ou la nature du suffixe de dérivation (s'il s'agit d'un nom dérivé).

2° Suffixes casuels. On distingue en mongol les cas suivants :

Génitif : après les voyelles et *y* : *-yi(n)*, *-gi(n)* ; après *n* : *-u*, *-i*, *-ai* ; après les consonnes : *-u(n)*, *-yi(n)*, *-ayi(n)*.

Accusatif : après voyelle : *-igi*, *-yi* ; après consonne : *-i*.

Datif-locatif : *-dur* (*-tur*), *-du* (*-tu*), *-da* (*-ta*), et *-a*.

Ablatif : *-atša*, *-ā*.

Instrumental : après voyelle : *-bar*, *-gar* ; après consonne : *-iyar*.

Coopératif : *-lu'a* (*lā*, *lō*), *lar*.

On peut y ajouter un cas limitatif *-čăgä* (*-čāy*) apparenté, semble-t-il, avec le comparatif turc en *-tšäg*, *-tšälik*, *-tšä* (sur les rapports entre les cas comparatif et limitatif, voir J. Deny, *Gramm. turque*, § 905, 903 et 1387).

De même qu'en turc, le substantif peut figurer dans la proposition sans aucune désinence casuelle. Tout comme en turc (voir p. 207), la chose se produit lorsque le substantif est : 1° sujet de la proposition ; 2° complément direct indéterminé ; 3° complément nominal dans une expression fixée par l'usage.

Le complément direct (déterminé) prend la forme de l'accusatif.

Le datif-locatif cumule les attributions du datif (allatif) et du locatif turcs. Le turc présente d'ailleurs des traces d'une confusion analogue (*Gramm. turque*, § 282).

L'ablatif et l'instrumental ont la même valeur que les cas correspondants en turc. Ajoutons que pour exprimer le degré comparatif en mongol, comme en turc, il suffit de mettre le terme de comparaison à l'ablatif.

Le mongol a, en plus, un cas coopératif, mais on notera que l'osmanli ancien a tenté de développer le cas opposé par le sens, sorte de cas de carence (avec *siz*) (*Gramm. turque*, § 881, remarque).

3° Suffixes possessifs. Cette catégorie grammaticale diffère de la catégorie correspondante du turc, non seulement par la place qu'elle occupe dans le mot, mais aussi quelque peu par sa nature.

Le mongol a des suffixes possessifs réfléchis qui s'ajoutent au substantif pour indiquer qu'il s'agit d'un objet appartenant à l'*auteur* de l'action dont cet objet est le complément. Ils prennent la forme du suffixe *-bän* (*-iyän* après consonnes), formes plus anciennes : *-yübän* ou *-yü'än*. Ces suffixes se joignent aux différentes désinences casuelles ; mais à l'accusatif et au génitif, ils peuvent s'y substituer purement et simplement. A l'ablatif et au datif-locatif, ils se combinent avec la désinence casuelle.

Ex. : Instrumental : *ältši-bär-iyän ilägäkü* (pron. moderne *ilăxu*) « envoyer par son ambassadeur ».

Quant au suffixe possessif tel qu'il existe en turc (non réfléchi),

il est représenté en mongol par le génitif des pronoms personnels.

Ex. *erdeni minu* (*tšinu*) « mon (ton) bijou »,

ügä manu « notre parole », formule archaïque placée au début des édits des princes vassaux (comparer le turc *söz-ümüz* des khans de Crimée).

Le seul fait que le génitif du pronom se place *après* le substantif, qu'il devrait précéder en sa qualité de complément (comparer p. 192), montre qu'il s'agit d'un élément tendant à devenir un suffixe (possessif). Il le devient même tout à fait dans la langue parlée, où *minu* aboutit à *min* ou *mi* après une voyelle longue, *-m* après une voyelle brève et *-im* après les consonnes.

Ex. : *gar-tä-m* « dans ma maison ».

On remarquera à première vue la ressemblance de ce suffixe avec le possessif turc -[*i*]*m* (tableau de la p. 210).

Cette curieuse concordance, jointe à la grande ressemblance des formes du pronom personnel dans les deux langues, permet de supposer que le turc a suivi une évolution analogue et que ses suffixes possessifs sont également partis du génitif des pronoms. Si elle se confirme, cette hypothèse constituera un sérieux argument en faveur de la communauté d'origine du turc et du mongol.

Le génitif du pron. de la 3e pers. *anu*, *inu* (cp. turc *aniñ*); abrégé dans la langue parlée en *-ni*, peut comme le turc -[*s*]*i* accompagner le 2e terme d'un « rapport d'annexion » (voir p. 208). Serait-ce l'origine de l'*n* pronominal turc ? Voir ci-dessus et p. 207.

Ex. : *xagan-u närä inu* (kalm. *xān-i närä-ni*) « le nom de l'empereur ».

4° Suffixe *-gi* ou *x'i*. Cette désinence correspond, jusque dans le détail, à la désinence turque *-ki* (p. 208) et, comme celle-ci, ne se joint qu'au locatif et au génitif.

Ex. *gadzar-da-gi* « (celui, celle) qui est dans (sur) la terre » (comparer turc *yär-dä-ki*, même sens).

II. *Copule mongole.*

Le mongol se sert du verbe « être » (*bügü*, *büxü*) plus volon-

tiers que le turc. Il l'exprime au moyen de plusieurs formes dont la plus usitée est *buyu*, forme mod. *biy*. Il existe aussi un verbe *bol-* correspondant au turc *bol-*, osm. *ol-* « devenir, être ».

A la différence du turc, les pronoms personnels (enclitiques) ne sont pas employés comme copule. Cette observation concerne le mongol littéraire et, sans doute, le mongol commun ; mais certains dialectes modernes ont développé la copule-pronom ou suffixe prédicatif.

En bouriate la copule est :

à la 1re pers. : sg. *-p* (*-m*), pl. *-bdi* (*-mdi*),

à la 2e pers. : sg. *-š*, pl. *-t*, *-ta*, *-tä* ;

à la 3e personne, on ne met rien (suffixe zéro).

Ces suffixes sont d'anciens pronoms enclitiques. Le parallélisme avec le turc est frappant. On remarquera aussi que le phénomène ne concerne que les deux premières personnes. Le pronom de la 3e personne n'est pas devenu copule (comparer le turc où il a cessé de l'être).

III. *Suffixes désinentiels du verbe mongol.*

1° Suffixes des formes nominales du verbe. Ils comprennent ce que M. Ramstedt appelle les « nomina verbalia » et les « converbia ».

Les « nomina verbalia » peuvent servir à former des quasi-propositions substantives (voir p. 212) aussi bien que des quasi-propositions adjectives. Ils réunissent donc les deux catégories, que nous avons pu distinguer plus nettement en turc, de formes substantives et de formes adjectives du verbe.

Les « converbia » sont des formes adverbiales du verbe et correspondent par conséquent aux gérondifs et locutions gérondives du turc (quasi-propositions adverbiales).

2° Formes verbales proprement dites ou formes prédicatives (que les grammairiens de l'école russe appellent aussi définitives, accomplies, achevées ou « verbum finitum »).

Certaines de ces formes n'apparaissent que dans cet emploi ; tels sont l'impératif, le prescriptif, le précatif, le bénédictif, etc. (d'après la terminologie de M. Ramstedt).

D'autres sont en réalité des formes nominales du verbe employées en fonction de forme prédicative, tout comme en turc.

On peut dire que les unes et les autres sont employées comme formes personnelles avec cette restriction pourtant que, en mongol commun, la personne n'est pas exprimée par un suffixe, mais par le pronom autonome ou le nom qui est sujet du verbe.

Ex. : *bi* (*tši*, *bidä*(*n*), *ta*) *gar-bay* « je (tu, nous, vous) suis... sorti » (prétérit).

Il est vrai que déjà dans la langue écrite on a souvent répété le pronom personnel en le faisant figurer après la forme verbale. Bobrovnikov (*Gram.*, p. 234) qui énumère les cas où ce fait se produit observe avec raison que le pronom ainsi répété a en réalité la valeur d'une désinence personnelle.

Le bouriate et les dialectes qui ont développé des copules-pronoms ont poussé l'évolution jusqu'à constituer des formes personnelles analogues à celles du turc.

Il résulte de ce qui précède que s'il y eut jamais un turco-mongol commun, cette langue ignorait les désinences personnelles, comme elle ignorait les suffixes possessifs. A part les pronoms eux-mêmes, toutes les formes pronominales indiquées pour le turc au tableau de la p. 210 semblent bien être des morphèmes tardifs, et leur absence en mongol ne doit pas être invoquée comme argument contre la théorie de l'unité d'origine.

Écriture.

Jusqu'au XIIIe siècle, les Mongols ne paraissent pas avoir éprouvé le besoin d'écrire leur langue.

Ce n'est qu'après 1204, c'est-à-dire après la défaite du prince des Naïman de la Mongolie occidentale, Tay yang Khan, qu'un certain Tatatonga, homme de confiance de ce dernier, aurait introduit chez les Mongols de son nouveau maître Gengiskhan l'écriture dite *ouïgoure*. Dans ses relations avec la dynastie des Kin, qu'il devait bientôt abattre, Gengiskhan se servit, il est vrai, en 1216, de caractères chinois, mais ceux-ci ne supplantèrent pas l'alphabet ouïgour qui fut désormais employé normalement.

Toutefois, à partir de 1269 et jusqu'au milieu du XIVᵉ siècle, on fit usage officiellement d'une écriture particulière adoptée par Khoubilaï-Khan et appelée *dürbäldẓin* « carrée », d'après sa forme, ou *'Phagspa*, d'après le nom (ou plutôt d'après l'épithète) du célèbre lama, son inventeur, originaire de Sa-skia (au Tibet). Cet alphabet dérivé de l'alphabet tibétain, comportait 41 caractères fondamentaux dont neuf pour figurer des voyelles. Il représentait aussi des sons aspirés et offrait l'avantage de noter avec une rare précision le timbre des voyelles. (Voir G. Pauthier, *Journ. as.*, janvier 1862, p. 1 à 47 ; pour les textes en écriture *'phags-pa*, voir Roland Bonaparte, *Documents de l'épigraphie mongole des XIIIᵉ et XIVᵉ siècles* et articles *Journ. as.*, 1894 [Chavannes, Sylvain Lévi, Hirth et Radloff] et 1896 [Deveria et Bang].)

L'écriture ouïgoure à laquelle on revint après l'abandon de l'alphabet 'phags-pa considéré, à juste titre d'ailleurs, comme compliqué et incommode, est en usage aujourd'hui. C'est cette écriture qu'on appelle mongole nationale, bien qu'il n'y ait entre elle et les types plus anciens de l'écriture ouïgoure qu'une différence de « ductus » ou de « main ». Elle a simplement pris, à la longue, un aspect plus anguleux, et les Mongols l'écrivent aujourd'hui *toujours* de haut en bas, en commençant par la gauche du papier.

Les Kalmouks ont perfectionné cet alphabet en 1648 (réforme de *Zaya Paṇḍiṭa*) en créant, par simple addition de points ou signes diacritiques, sept caractères nouveaux qui permettent d'éviter les incertitudes inhérentes à l'usage de l'alphabet ouïgour et de figurer beaucoup plus exactement la prononciation de la langue parlée.

Les Mongols proprement dits et les Kalmouks ont de plus inventé un certain nombre de lettres, dont quelques-unes servent uniquement à transcrire les mots sanskrits (alphabet galikh) et d'autres les mots tibétains.

Le plus ancien monument écrit en mongol actuellement connu est une inscription de cinq lignes, contemporaine de Gengiskhan, gravée sur granit et où il est fait mention de son neveu *Yisunkä*. Trouvée sur l'Onon près de Nertchinsk, conservée au Musée asia-

tique de Pétrograd, elle a été traduite et publiée par J. J. Schmidt, par Banzarov et par Pozdnieev.

Vient ensuite le texte d'un cachet mongol apposé en 1246 sur une lettre du grand khan Guyuk au pape Innocent IV, et que rapporta Plan Carpin (conservée à la Vaticane). Déchiffré par M. Pelliot, ce texte doit être publié prochainement par lui.

Il faut placer ensuite une série d'inscriptions en écriture 'phags-pa, les lettres des ilkhans de Perse Argoun et Euldjaïtou à Philippe le Bel, publiées par Abel-Rémusat et trois nouvelles lettres des ilkhans aux papes que M. Pelliot éditera prochainement.

Littérature.

La littérature mongole est relativement peu connue. La partie la plus importante et la plus ancienne en est constituée par les textes historiques, notamment *l'Histoire secrète des Mongols* (en chinois : Yuan tch'ao piche), en 75 chapitres, écrite dès 1240, perdue en caractères mongols, mais conservée intégralement dans une transcription phonétique en écriture chinoise qui permet de rétablir le texte mongol original.

Il existe aussi un grand nombre de traductions de textes bouddhiques, quelques textes chamanistes, des livres didactiques, quelques livres juridiques et médicaux, sans compter les chants populaires (épiques et autres).

Dans l'ensemble, littérature peu originale, voir : Laufer, *Skizze der Mongolischen Literatur*, Keleti Szemle 1907 (pp. 164-261).

Les Mongols se sont servis de l'imprimerie dès le commencement du XIVe siècle (traduction du Kandjour tibétain qui fut reprise et terminée en 1623 ; M. Pelliot a rapporté une poésie mongole non bouddhique imprimée au XIVe siècle).

La disparition de la dynastie des Yuan ne ralentit pas les études mongoles en Chine. En 1407, l'empereur Yong-Lo fait réorganiser l'École des Interprètes ; le mongol y forme la première section (sur les huit constituées) avec 38 directeurs ou professeurs. Cette école fut reconnue en 1644 (sous les empereurs mandchoux) et le mongol y conserva la première place.

En Corée, le mongol occupa jusqu'en 1469 la deuxième section de l'Ecole des Interprètes (après le chinois et avant le japonais et le niu-tche).

BIBLIOGRAPHIE

Il existe peu d'ouvrages en français sur la langue mongole. Comme manuels, on peut citer : Georges Soulié, *Gramm.* (dial. ordos), Paris, 1903 ; Vitale et de Sercey, *Gramm. et voc.* (dial. des Khalkhas, Péking, 1897 ; Feer, *Table de la langue mong.* Paris, 1816 (7 folios autographiés).

Il faut y ajouter la traduction (lithographiée à Rennes, 1870) de la grammaire de Schmidt (voir plus loin).

En outre l'introuvable dictionnaire de Kovalevski contient une partie française.

En anglais on a encore moins d'ouvrages : Juille, *Gramm.* (xylographiée 1838 ; voir Eitel, *Handbook of chinese Buddhism* (termes mongols), Londres, 1888, voir aussi Stanley Lane-Poole, *Catal. of Or. Coins*, t. VI (numismatique mongole).

En allemand on trouve les travaux de Gabelentz (*Mandschu mong. Gramm.*, Göttingen, 1837) de I. J. Schmidt (*Gramm.*, Petrograd, 1831, traduite en fr., en russe, Pet. 1832 et adaptée en italien par Puini, Florence, 1878 ; Dictionn. mongol-allemand-russe, Pét., 1831), de Castrén (Gramm. bouriate, Pét., 1857), de Bang (articles) et surtout de Ramstedt (professeur à Helsingfors).

En hongrois, études de Budenz et de Balint.

C'est en russe qu'ont paru les travaux les plus importants : Bobrovnikov est l'auteur d'une grammaire du Kalmouk (Kasan, 1849) qui est encore aujourd'hui la meilleure grammaire mongole qu'on possède. Indépendamment du dictionnaire de Kovalevski (Kasan, 1846-1848 en 3 vol.), on a celui de Golstounski (1893-1895), 3 vol., complété en 1901 avec la collaboration de Roudnev. Voir aussi les grammaires, chrestomathies et vocabulaires de Popov, Orlov, Podgorbounski, etc. et les travaux de Kotvich, Mélioranski et surtout Pozdnieev et Roudnev. Grâce à ces deux derniers auxquels il faut joindre Ramstedt, les études mongoles ont reçu une impulsion énergique.

Il existe enfin un article de M. V. Thomsen sur la langue mongole (en danois) dans *Salm. Konv. Lex.*, t. XII, p. 942, 1091.

Pour plus de détails voir l'article de M. Laufer cité plus haut (p. 232) et les bibliographies placées en tête de l'ouvrage de Roudnev, *Materyali po govoru vost. Mong.* Pétr., 1911 (en russe), et à la fin de *Xori-buryatskiy govor*, Pétr., 1913-1914 (en russe), du même.

LANGUES TONGOUZES

Les Tongouzes et Mandchoux.

Certains auteurs font remonter l'histoire des Tongouzes à l'apparition des Sou-Chen (ou Si-Chen), nomades qui auraient apporté en 2225 av. J.-C. (?) un tribut d'arcs et de flèches à Chouen, légendaire empereur de Chine. « Tongouze » vient de *donki* « hommes ». Cette peuplade habitait encore au IIIe siècle av. J.-C. en Mandchourie actuelle, dans la région de Kirin (ou Guirin) sur le cours supérieur et moyen du Soungari, affluent de droite de l'Amour. Changeant ensuite de nom suivant le clan exerçant l'hégémonie, les Sou-Chen seraient devenus successivement les Yi-leou sous la dynastie des Hans, les Wou-ki sous les Weï (IIIe s. ap. J.-C.) et les Mo-ho (sur le cours de l'Amour) sous les T'ang. Une branche des Mo-ho, les Sou-mo Mo-ho fondèrent au VIIe siècle le royaume civilisé de Po-haï (nom de la baie de Corée) qui s'étendait sur toute la Mandchourie, l'Oussouri inférieur et la Corée.

Au milieu du Xe siècle ce royaume est renversé par les K'itan (voir p. 220).

En 1125 les K'itan sont renversés eux-mêmes par la tribu tongouze des Niu-tche qui s'appelaient avant 1032 les Niu-tchen et qui étaient d'anciens tributaires de la Corée. Les Niu-tche connus aussi sous le nom de *Džurdžin* ou Jou-tchen, seraient les descendants des Sou-chen par les Mo-ho nommés plus haut. Leur maison royale s'appelait *kin* en chinois, *altšun* en leur langue d'où mandchou *aisin,* ce qui signifie « (dynastie d)'or » (comparer le turc *altun* « or »). Ils régnèrent sur la moitié septentrionale de la Chine et furent gens civilisés.

En 1234, les Mongols abattent la puissance des Kin (voir p. 220) et s'annexent la Mandchourie. Un siècle et demi plus tard, les Mongols sont chassés de ce pays où cherchent à pénétrer, avec plus ou moins de succès, les Ming. Parmi les tribus tongouzes sévit à ce moment la guerre intestine.

Le chef *Nurxatsi* réunit alors les sept *ayman* (synonyme mandchou du mongol *aymaq*) des Niu-tche et se fait proclamer empereur en 1616 ; il installe sa capitale à Moukden en 1625. T'ai-tsou, tel est le nom de temple ou posthume (miao-hao) de ce prince, est l'ancêtre de la dynastie des Mandchoux (*Mandẓu*), qui, après avoir renversé la dynastie des Ming en 1644, commence à régner en Chine dans la personne du fils de *Nurxatsi*, *Abaxay* ou T'ai-tsong. En 1636, ce dernier donne à sa dynastie le nom chinois de Ta Ts'ing ou Ts'ing sous lequel elle fut connue jusqu'à l'abdication du dernier empereur, le tout jeune P'ou-Yi, de son nom de règne Sinan-t'ong, c'est-à-dire jusqu'à la proclamation de la République (12 février 1912).

On ignore l'origine du nom des Mandchoux qui n'apparaît qu'au commencement du XVIIe siècle pour désigner l'ensemble des nouveaux conquérants de la Chine. Quant à la « Mandchourie » elle n'est ainsi appelée que par les Européens. Les Chinois l'appellent Cheng-king, et surtout Tong-san-cheng, les « trois provinces orientales ».

Comme nous l'avons déjà indiqué (p. 219), l'origine tongouze des Sienpi et des K'i-tan est actuellement contestée. Il en est de même des Weï « postérieurs » (Yuan Weï) ou T'o-pa qui régnèrent dans la Chine du Nord de 386 à 535 et pour la langue desquels le *Nants'i chou* qui porte sur la fin du Ve siècle et a été rédigé au début du VIe, donne quelques mots en transcription. Ces mots ne sont pas tongouzes, mais on n'a pu encore déterminer s'ils sont spécifiquement turcs plutôt que mongols (voir Pelliot, *L'origine de T'ou-Kiue...*, T'oung-Pao, 1915, p. 689).

Distribution géographique. — Les Tongouzes habitent à l'Est de l'Iénisséï sur le cours des deux Toungouzka. Ensuite, comme voisins méridionaux des Yakoutes et orientaux des Mongols (Bouriates et autres), ils s'étendent au Sud de la Lena, jusqu'à la

mer d'Okhotsk (moins la presqu'île du Kamtchatka) et la mer du Japon (moins la Corée).

Faute de renseignements il est impossible de donner une classification exacte des dialectes tongouzes.

D'après Schrenck, on peut distribuer ainsi les peuples qui les parlent (pour le bassin de l'Amour seulement).

1) rameau du Sud ou mandchou :

a) *dahur*, *solon* (fortement mélangé de Mongols).

b) mandchou, *gold*, *oroš*.

2) rameau du Nord ou de Sibérie :

a) *orotšon*, *manegir*, *birar*, *kile* (sur la riv. Kour).

b) *oltša* (sur l'Amour), *oroq* (l'île de Sakhaline), *negda*, *samagir*.

Sternberg qui critique cette classification estime qu'on ne saurait, provisoirement, envisager qu'une classification géographique.

A ce point de vue, on peut distribuer les populations tongouzes de la façon suivante :

1) Mandchoux, en Mandchourie (carte : 1), avec leurs voisins immédiats : *Dahur* (2), *Solon* (3), *Manegir* (4), *Birar* (5), *Gold* (6).

2) diverses petites peuplades de la Sibérie orientale (à l'Est et au Nord de la Mandchourie) : *Oroš* (carte 7), *Kile* (8), *Oltša* (9), *Samagir* (10), *Negda* (11), *Inkagir* (12), *Lalegir* (13), *Lamut* (15), *Utšur* (16), *Maya* (17), *Bital* (18), *Kangalas* (19), *Orotšon* (20), *Oroq* de Sakhaline (23). — Les Mandchoux appellent *orotšon* (« chasseurs de rennes ») tous les Tongouzes en général. Certains appellent « Tongouzes proprement dits » les *Orotšon* et les *Lamut* (ou « Maritimes »). Les *Oltša* s'appellent aussi *Mangun*.

3) Tongouzes de la région de l'Iénisséï ; (carte : 21) et *Tsapogir* (22).

4) *Šibo* (*Šibä* ou *Sibo*) et *Solon* dans la vallée de l'Ili à la frontière russo-chinoise.

Les Mandchoux qui occupent les deux rives du Khourkha et la rive droite du Soungari sont cultivateurs et éleveurs ; les autres Tongouzes sont très peu civilisés, beaucoup sont nomades,

d'autres semi-sédentaires et chasseurs de rennes. Certains ont abandonné la chasse pour la pêche.

Le recensement chinois de 1918 accuse 1.380.440 Mandchoux adultes, dans les huit « bannières » qu'ils forment administrativement en Mandchourie (l'ensemble de la population de cette province y compris les Chinois, les Mongols et d'autres est de 12 millions et demi). Rockhill estime qu'en ajoutant à ce nombre les enfants et les garnisons mandchoues des villes de Nan-King (prov. de Kiang-Sou), King-tcheou (Hou-pe) et Hang-Tcheou (Tche-Kiang), on obtiendrait un chiffre global de 1.500.000 Mandchoux en Chine.

D'autre part, d'après le recensement russe de 1897, il existe en Sibérie 3.394 Mandchoux (sur la rive gauche de l'Amour, en face d'Aïgoun) et *seulement* 73.110 Tongouzes pour l'énorme espace qu'ils occupent sur le territoire russe, ce qui représente une densité vraiment infinitésimale. Voici d'ailleurs leur distribution par provinces (avec l'indication des chiffres de 1897 et de ceux de 1911) : pr. d'Amour — 5.005 > 1.328 ; pr. Primorskaïa (du Littoral) — 9.298 > 7.896 ; de Yakoutsk — 12.231 > 12.930 ; d'Irkoutsk — 2.192 > 2.614 ; de Transbaïkalie — 34.403 > 39.593 ; d'Iénisséi — 3.172 > 3.276 ; de Tobolsk — 4 ; de Tomsk — 3 ; de Kamtchatka — 9.311 > 7.216 ; l'île de Sakhaline — 892 > 351. Un seul Tongouze est signalé comme domicilié en Russie d'Europe (à Moscou). Le nombre total des Tongouzes de Russie passe de 76.505 en 1897 à 75.204 en 1911.

Comme la plupart des Mandchoux ont abandonné leur langue pour le chinois (voir plus loin p. 239 les points où ils ont conservé leur langage maternel), on est amené à constater que l'ensemble de la population de langue tongouze-mandchoue est loin d'atteindre un million d'individus.

On a vu que le nombre des Tongouzes décroît. Dans le district d'Okhotsk ils ont diminué de moitié en 131 ans. Schrenck attribue cette décroissance à la dureté du climat, à l'insouciance, à la paresse, au mépris de la femme, aux attaques des cosaques russes et à l'oppression exercée par les fonctionnaires mandchoux et chinois. Il serait donc nécessaire de relever le plus tôt possible les parlers de ces peuplades appelées à disparaître.

Écriture.

A l'imitation de leurs prédécésseurs (les K'itan) qui en 920 avaient adapté à leur langue les hiéroglyphes chinois, les Niu-tche inventèrent un alphabet du même genre en 1119. Un deuxième alphabet fut créé peu après sous le nom de « grand alphabet » par opposition au premier qui reçut celui de « petit ». Cette écriture a été déchiffrée par Wilhelm Grube (*Note prélim. sur la langue et l'écrit. des Jou-tchen*, T'oung Pao, 1894, p. 334-340). Les Niu-tche soumis aux Mongols adoptèrent l'alphabet ouïgour, mais leur alphabet national se maintint encore en Mandchourie jusqu'en 1526. Les documents rédigés en cette écriture ne nous sont parvenus qu'en très petit nombre.

Comme l'écriture jou-tchen était extrêmement compliquée (plus de mille signes), les Mandchoux adoptèrent l'alphabet mongol (ouïgour) dès le règne de Nourhatsi en 1599. En 1632, on ajouta quelques signes diacritiques. A part ces signes, il n'y a qu'une légère différence de « main » entre l'écriture mandchoue et l'écriture mongole. Par imitation des habitudes de chancellerie chinoises, l'empereur K'ien-long donna l'ordre, en 1848, d'élaborer, indépendamment des trois modèles d'écriture — l'officielle, la semi-officielle et la cursive — 32 variantes d'écriture mandchoue carrée pour faire pendant aux 32 variantes également artificielles de l'ancienne écriture chinoise. Quelques-unes de ces variantes furent employées pour graver les cachets officiels.

Les imprimés mandchoux les plus anciens connus remontent à 1647. Ils ressemblent extérieurement aux livres chinois et s'impriment non avec des caractères mobiles, comme ce fut souvent le cas en Corée, mais en xylographes. Ils sont paginés de gauche à droite et non de droite à gauche comme les livres chinois.

Le mandchou langue écrite.

Ce qui frappe dans le domaine des langues tongouzes, c'est la disproportion entre la place qu'y tiennent les dialectes parlés et celle qu'y occupe la langue mandchoue écrite. Celle-ci eut la

bonne fortune d'être conservée, développée et codifiée par une série d'empereurs intelligents. Rien de plus instructif que la sollicitude avec laquelle les empereurs mandchoux, surtout K'ang-Hi (1662-1722) et K'ien-long (1736-1795), défendent cette langue contre l'invasion des mots chinois, font rédiger des manuels, traduire des ouvrages chinois et mongols, établir de gigantesques dictionnaires polyglottes. Le mongol n'était pas d'ailleurs négligé. En 1708 paraît un « Miroir des langues mandchoue et mongole » en 21 volumes, en 1780 un ouvrage analogue en 24 volumes, sans compter un dictionnaire mandchou-mongol-chinois-tibétain en 10 volumes, un autre en 36 volumes. Le British Museum possède un exemplaire manuscrit unique, en 36 volumes, d'un dictionnaire en cinq langues (la 5e étant le turc tchaghataï). En 1780 paraît un aperçu littéraire du mandchou, avec grammaire, en 12 volumes, qui a été utilisé en partie par Gabelentz. Il serait trop long d'énumérer les autres ouvrages de cette nature.

Les fonctionnaires étaient tenus à un examen de mandchou.

Malgré tous ces efforts, la langue mandchoue s'altérait de plus en plus — comme prononciation et comme grammaire — au contact du chinois et ne vivait que d'une vie artificielle dans la capitale. Après la mort de l'empereur Hien-fong (1761), les examens de mandchou sont abandonnés et cette langue n'est plus connue à Péking que de quelques isolés.

L'histoire du mandchou nous offre donc un curieux exemple de sauvetage officiel d'une langue écrite destinée à disparaître aussi bien à cause de la supériorité de la culture chinoise que du petit nombre des Mandchoux.

Il ne faut pas oublier cependant, comme on l'a fait pendant ongtemps, que le mandchou, ainsi que l'a établi M. Roudnev, est aussi une langue vivante parlée dans le pays de l'Ili où il y eut des garnisons mandchoues et dans la Mandchourie du Nord, c'est-à-dire dans la province de Hei-long-Kiang (dans les villes d'Aïgoun, Merguen, Boutkha, Tsitsikar et les villages environnants) et dans celle de Kirin ou Guirin (dans les districts de San-si, Ningouka, Ache-ho et Kirin).

Le mandchou, en tant que langue morte, a été pendant longtemps considéré comme un auxiliaire important pour l'interprétation des textes chinois à cause des traductions interlinéaires en langue mandchoue plus facile à apprendre.

De quelques particularités phonétiques.

Le mandchou possède une voyelle longue (*ō*).

La loi de l'harmonie vocalique peut, malgré ses fluctuations, se résumer ainsi en mandchou :

voyelles postérieures : *a*, *o*, *ō*. Ex. : *tsaqōran* « arbre ».

» antérieures : *ä*, *u*. Ex. : *ämu* « un ».

» indifférente : *i*. Ex. *miñgan* « mille », *däelixun* « rate ».

En tongouze :

voyelles postérieures : *a*, *o*, *u* ; Ex. *aruqun* « lentement » (cf. osm. anc. *arqun*).

» antérieures : *ä*, *ö*, *ü*, Ex. *mäñün* « monnaie, argent ».

» indifférente : *e*, *i* *i̥*, Ex. *ätirkän* « vieillard », *atirqan* « vieille femme ».

La voyelle *i* exerce en mandchou une grande influence sur les consonnes qui la précèdent. Devant *i*, *k* et *tš* deviennent *ts* : *g* et *dž* deviennent *dz* ; *s* devient *š*.

Lorsqu'un mot est terminé par *n*, il perd cette consonne en recevant un suffixe qui commence par une consonne :

ahôn « frère aîné », pl. *ahôta*.

Il en est de même parfois d'une voyelle finale : *ilha* « plante » pour *ila-ha* participe passé de *ila* « fleurir ».

Les crases de deux mots paraissent relativement plus fréquentes en mandchou et tongouze qu'en turc et mongol.

memama pour *meme ama* « père nourricier ».

tsihaqô pour *tsiha* + *aqô* « involontairement, « proprement » sans volonté ».

Dans la langue écrite, ce procédé de crase a été en outre développé abusivement pour fournir, d'une manière artificielle, des termes mandchoux correspondant à des termes chinois dont le mandchou n'avait pas l'équivalent.

Suffixes désinentiels du tongouze et mandchou.

Malgré l'importance que conserve en mandchou et tongouze la formation suffixale il est bon de rappeler que ces langues accusent un état flexionnel sensiblement plus marqué que le mongol.

I. Suffixes désinentiels des noms tongouzes et mandchoux.

1° Suffixes du pluriel. — En mandchou, la marque du pluriel ne s'emploie que pour un petit nombre de mots désignant presque exclusivement des personnes [suffixes : *sa* (*so*, *sä*), *ta* (*tä*), *si*, *ri*, *säi*].

mandžu « mandchou », pl. *mandžu-sa*.

En tongouze, le pluriel est plus souvent employé. Suffixes : *-l*, *-r*, *-il*, *-ul*, *-sal*, *-ral*, *-nasal*.

bira « rivière » ; plur. *bira-l*.

Le mandchou, par imitation des « numéraux » du chinois, s'est donné un certain nombre de pluriels de classes (*da* pour les arbres ; *fulga* maisons ; *fali* êtres inanimés etc.).

2° Déclinaison. — Il n'existe pas à proprement parler de suffixes casuels en mandchou. Les divers cas sont exprimés au moyen de postpositions (n'obéissant pas à la loi d'harmonie vocalique). En tongouze les particules correspondantes sont devenues de véritables suffixes.

Le mandchou possède les cas suivants :

Accusatif : *bä* (tongouze *-vä*).

Génitif instrumental : *i*, *ni* (tong. *-ñi*).

Datif-locatif : *dä* (tong. *-du*).

Ablatif : *tsi* (tong. *-ḍuq*).

De même qu'en turc et en mongol, le substantif nu peut exprimer le complément direct et le complément nominal.

La déclinaison jou-tchen a la même forme pour le génitif, une forme analogue pour le locatif (*to*, *tu*) et le suffixe *-roh* à l'accusatif.

Le tongouze a un grand nombre de cas. Il possède en plus

des cas du mandchou : le prolatif (suff. *-duli*, *-li*), le comitatif (*-nun*) et l'instrumental (*-dži*, comparer l'ablatif du mandchou).

3° Rapport possessif. — Il n'existe pas de suffixe possessif en mandchou ; on y emploie pour exprimer un rapport possessif pronominal le génitif du pronom personnel qui se met avant le nom possédé, c'est-à-dire à la place normale du complément nominal (contrairement à l'usage mongol).

Ex. : *mini* (*sini*, *mäni*, *suwäni*) *džusä* « mon (ton, notre, votre) fils ».

L'objet possédé (*džusä*) ne recevant aucun suffixe et l'objet possédant pouvant ne pas recevoir la marque du génitif (voir plus haut), il en résulte qu'un rapport possessif peut être représenté en mandchou par la simple juxtaposition de deux noms, en mettant le complément en tête : *boo kooli* « les rites de la maison ».

Le tongouze use du même procédé, mais il se sert également des *suffixes possessifs* qui sont à la 1re et 2e personne d'anciens pronoms personnels suffixés ou enclitiques (comme en turc et en mongol).

Pronoms personnels		Suffixes possessifs.
1re pers.	sg. *bi*	*-f*, *-u*
	pl. *bu*	*-vun*
2e pers.	sg. *ši*	*-s*
	pl. *su*	*-sun*, *-hun*

II. *Copule tongouze et mandchoue.*

Elle est représentée en mandchou et en tongouze par le verbe, *bi* « être » ; mandchou *bimbi* (comp. tongouze de Nertchinsk : *bi-ši-m*, *bi-hi-m*) « je suis, tu es... » ; *bimbi he* « j'étais ».

III. *Suffixes désinentiels du verbe tongouze et mandchou.*

1° Suffixes des formes nominales du verbe.

Il existe en mandchou et en tongouze d'assez nombreuses formes substantives (infinitif), adjectives (participes) et adverbiales (gérondifs). Certains gérondifs sont constitués au moyen d'infinitifs mis à des cas appropriés de la déclinaison.

2° Formes verbales proprement dites ou formes prédicatives (ou définies, comme dit Adam).

Le mandchou se comporte ici comme le mongol, c'est-à-dire qu'il fait précéder tantôt des formes personnelles spéciales tantôt des participes du pronom personnel (autonome).

Ex. *bi* (*bä*, *si*, *suwä*) *ara-mbi* « je (tu, il, nous, vous) écris... » (indicatif présent).

Le tongouze possède des formes personnelles avec suffixe (personnel) prédicatif.

Ces suffixes prédicatifs sont des suffixes possessifs (= conjugaison subjective voir plus haut, p. 213), sauf au présent de l'indicatif où il y a un suffixe prédicatif particulier, également d'origine pronominale.

BIBLIOGRAPHIE

Manuels en français : Hans Conon von der Gabelentz, *Éléments de la grammaire mandchoue*, Altenbourg, 1832 ; Lucien Adam, *Grammaire de la langue mandchoue*, Paris, 1873 ; le même, *Gr. de la langue tongouze*, 1874 ; Harlez, *Manuel de la langue mandchoue*, 1884. — Travaux divers d'Amyot (dictionnaire), Langlès, Abel-Rémusat, Harlez, Devéria.

En latin : *Elementa linguae tartaricae*, sans nom d'auteur, publiés par Thevenot; traduction française, Paris, 1787.

En anglais : Möllendorf, *A Manchu Grammar*, 1892. — Voir surtout les travaux de Wylie et Terrien de Lacouperie.

En allemand : A. Costien, *Grundzüge einer tung. Sprachlehre*, Saint-Pét., 1857, Gabelentz, *Mantschu-deutsches Wörterbuch*, Leipzig, 1864. — Travaux de Klaproth, Grube, Bang, Czekanoswki, Maydell, P. Schmidt, E. von Zach.

En italien : Travaux (récents) de Teza.

En russe : Orlov, *Gramm. mandchou*, Saint-Pét. 1873, Orlov, *Chrestomathie* 1863 ; Vassiliev, *Dict. mandchou-russe*, Saint-Pét., 1866 ; Zakharov, *Gramm.* 1879 ; du même, *Dict. manchou-russe* ; Ivanovski, *Chrestomathie*, Saint-Pét., 1771, travaux d'Ivanovski, Roudnev, Dobrovolski.

Pour plus de détails voir Laufer (Berthold), *Skizze der manjuischen Literatur*, Keleti Szemle, 1908, p. 1-53, et Grebenstchikof, *Mandžurį, ix yazįk i pis'm'ennost'* (en russe), dans le bulletin de l'Institut oriental de Vladivostok, t. XXXII, nº 2 (voir les notes au bas des pages).

J. DENY.

LA LANGUE JAPONAISE

Le nombre des études parues jusqu'à ce jour ne doit pas faire illusion sur la valeur des résultats acquis dans le domaine de la linguistique japonaise; ces travaux se ressentent de l'insuffisance de la préparation linguistique de leurs auteurs et nous donnent souvent des théories sur la langue japonaise, sans se tenir aux faits de la langue même.

Le ministère de l'Instruction publique du Japon recueille des matériaux sur les dialectes du Japon, mais ce grand travail n'est pas encore près d'être achevé. La langue japonaise moderne n'est pas encore suffisamment étudiée du point de vue phonétique : il nous manque une histoire des sons japonais; la morphologie, la sémantique sont peu connues dans leur développement historique, il en est de même pour la syntaxe et la stylistique ; bref une histoire complète de la langue japonaise reste encore à écrire, et c'est pourquoi nous ne sommes capables d'établir ni la parenté de cette langue avec d'autres, ni ses origines.

Les études déjà faites nous montrent qu'elle est apparentée à la langue des îles de Riou-Kiou et qu'elle a quelques points de contact avec la langue coréenne ; quelques emprunts à la langue des Aïnou subsistent également dans son vocabulaire.

*
* *

Par sa structure le japonais (dénomination indigène : *nihongo* ou *kokugo*) est du type dit « agglutinant » ; il accuse cependant une tendance plus marquée que le turc à unir étroitement la

racine avec les autres éléments, et cela surtout si on envisage la conjugaison des verbes.

Les éléments de la langue japonaise peuvent être divisés en éléments dépourvus de flexion et éléments flexionnels. Parmi les premiers éléments nous avons le nom, les pronoms et les adverbes.

En japonais le nom est invariable et ne distingue par sa forme ni le genre, ni le nombre de l'être ou de l'objet qu'il indique ; les relations syntaxiques sont marquées par des particules. Par exemple : *fude* « pinceau », *fude no saki* « le bout du pinceau », *fude de kaku* « écrire au moyen du pinceau », *fude wo arau* « laver le pinceau », etc. Le mot *fude* peut signifier « pinceau » ou « pinceaux » ; le pluriel est indiqué par un suffixe *nado* ; dans une phrase on peut dire *mina* qui veut dire « tous ». On indique le pluriel des personnes par des suffixes ; ainsi on dit *šosei* « famulus », *šoseidomo* « famuli », *sensei* « maître », *senseitatši* « maîtres » ; mais l'idée du pluriel est moins nette que dans les langues indo-européennes, et les mots comme *kodomo* « les enfants » et *tomodatši* « amis » bien que pluriels en raison de leurs suffixes n'en sont pas moins employés sous la même forme au singulier. Il y a aussi un pluriel marqué par le redoublement : *hito* « homme », *hitobito* « hommes ».

Les pronoms proprement dits sont peu nombreux et, étant donné que le verbe est impersonnel, sont beaucoup moins employés que dans les langues indo-européennes. Il y a d'un autre côté en japonais beaucoup de noms qui en raison de leur sens sont employés pronominalement.

Les adverbes japonais présentent une certaine richesse : on les forme de plusieurs façons :

1) en ajoutant *ku* aux formes réduites des adjectifs : *mižikai* « bref », *mižikaku* « brièvement » ;

2) en substituant la particule *ni* à la particule *na* des adjectifs d'origine chinoise : *bureina* « impoli », *bureini* « impoliment » ;

3) au moyen d'onomatopées avec la particule *to* : *sorosoro to* « doucement » ;

4) au moyen de la particule *to* devenue inséparable : *zutto* « droitement », *kitto* « certainement » ;

5) par le redoublement : *tabi* « fois », *tabitabi* « souvent », *tsuki* « mois », *tsukizuki* « mensuellement » ;

6) avec le gérondif employé comme adverbe : *hažimete* « d'abord », de *hažimeru* « commencer ».

Les éléments flexionnels sont le verbe, et l'adjectif dont nous parlerons plus loin.

Le verbe est impersonnel et indique un fait. Les flexions du verbe marquent les circonstances de *temps* et de *mode*. Certains verbes ainsi que certaines formes du verbe japonais sont censés moins polis que les autres et ne s'emploient pas à la 2ᵉ personne, d'autres au contraire sont des verbes honorifiques et s'emploient seulement pour la 2ᵉ et parfois pour la 3ᵉ personne.

Les verbes de la langue japonaise moderne ont trois conjugaisons régulières avec : un thème à consonne finale du type *yom*, « lire », *kak* « écrire », *sas* « piquer », un thème vocalique en *-e*, *-ake* « ouvrir », et un autre en *-i*, *mi* « voir », *otši* « tomber » — et en plus des verbes irréguliers *suru* « faire » et *kuru* « venir ». Dans les formes antérieures de la langue, les conjugaisons du verbe japonais étaient plus nombreuses.

Nous ne pouvons pas décrire toutes les formes des conjugaisons du verbe japonais, mais nous tenons à indiquer que la forme du futur en *ō*, comme *ikimasō* « est-ce que nous n'irons pas », est toujours une espèce de mode éventuel ; c'est le présent qui est employé pour le futur. Les formes négatives du verbe japonais moderne sont surtout composées du verbe + la négation *nai* qui est un adjectif, et pour former les différents temps on se sert du verbe + *nai* + le verbe « être » *aru*. Ainsi le passé « je n'ai pas lu » est *yom* + *a* + *nak* + *atta*.

Le verbe japonais a les voix suivantes : une forme passive *utareru* « être battu », qui a aussi un sens actif de « pouvoir battre », puis une forme causative *utaseru* « faire battre », et même

une forme passive causative *utaresaseru* « se faire battre ». De plus le verbe japonais a une forme désidérative en *-tai* : *kakitai* « vouloir écrire » et une forme « de vraisemblance » en *rašii* : *kaku rašii* « (il) paraît écrire ».

Le verbe japonais est toujours composé dans le même ordre : la racine + le thème + la voix + le temps + la négation ; si la négation est *naku* + verbe *aru*, les deux dernières parties changent de place et nous avons la négation + le temps : *yom a re nak atta* « cela n'était pas possible de lire ».

L'adjectif japonais a deux formes : une à flexion, et dans ce cas c'est plutôt un verbe intransitif, et une forme sans flexion où il affecte la forme nominale. Les adjectifs à flexion se présentent sous une forme telle que la suivante : soit *fuka* « profondeur » (qui ne s'emploie pas isolément dans la langue japonaise moderne) ; on a : *fuka* + *sa* « profondeur », mais *fuka* + *i* « profond » : *fuka* + *ku* « profondément » est un adverbe, et *fukaku* avec les différents temps du verbe « être » *aru* donne *fukakatta* « c'était profond », *fukakereba* « si c'est profond », etc. Les adjectifs sans flexion sont pour la plupart d'origine chinoise ; ils étaient originairement employés avec le verbe *naru* « être, devenir » ; *bimbō naru* « être pauvre » puis à la fin du XIIe siècle le verbe *naru* a perdu sa 2e syllabe ; il a subsisté sous la forme *na*, et il est résulté de là une classe d'adjectifs qualificatifs comme *kirei na* « beau », *burei na* « impoli », etc.

L'ordre des mots est relativement fixe en japonais ; la position des participes et des verbes principaux ne change pas, et en général la phrase ne présente pas d'inversion, ni d'interrogation indirecte. L'enchaînement des différentes parties est marqué par les particules qui ont diverses fonctions et qui indiquent les différents rôles des autres mots en devenant aussi prépositions et conjonctions. L'ordre est en général le suivant : sujet, complément, verbe. Le verbe est toujours à la fin. Par exemple :

Gakunen wa kugatsu žūitši nitši ni bažimatta. Sanširō ga
L'année scolaire septembre onze a commencé Sanshirō

学年は 九月十一に はじまった。

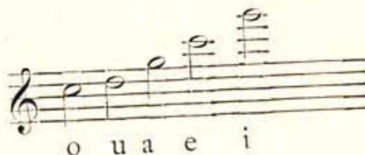
ẓitto šite ike no omo wo mitsumete iru to
immobile étang surface fixant (du regard) éta orsque
ōkina ki ga ikuhon to naku mizu no
des grands arbres beaucoup (sans quantité de troncs) l'eau
soko ni utsutte sono mata soko ni aoi
fond dans se reflétant (dans) ce de nouveau dans le fond bleu
sora ga mieru.
ciel était visible.

(Extrait de Sanshirō, roman de Natsume Sōseki).

*
* *

Si nous considérons le japonais du point de vue phonétique nous remarquons que ses sons n'offrent aucune difficulté d'émission ; il évite en général les groupes difficiles soit en intercalant des sons transitoires, qui deviennent indépendants, soit en remplaçant une combinaison difficile de deux consonnes par une consonne longue.

Le système des voyelles japonaises est très simple : trois voyelles antérieures *i*, *e*, *a*, — une voyelle postérieure, *o* — et une voyelle postérieure sansarrondissement, *u*. Pour prononcer la voyelle *u*, les lèvres restent dans une position neutre en formant une fente étroite, la langue touche presque les gencives inférieures et la pointe de la langue est légèrement recourbée vers le palais. Il y a une élévation de l'arrière-langue (Edwards, p. 17). De plus dans leur hauteur relative les voyelles japonaises occupent la place suivante :



La hauteur de l'*u* est plus grande que celle de l'*o*. Mais la voyelle *u* et la voyelle *i* ont une sonorité plus faible que les trois autres voyelles japonaises. Une voyelle japonaise, qu'elle

soit longue ou brève, a toujours le même son et garde sa netteté en syllabe faible ; en général les voyelles sont beaucoup plus stables que les consonnes. Grâce à son accent uni le japonais ne connaît pas de voyelle neutre.

Les consonnes japonaises se divisent en trois classes : les occlusives, les nasales et les fricatives.

Les occlusives *p* et *b* exigent une tension musculaire des lèvres et une fermeture assez faible. Dans la langue moderne le *p* initial existe seulement dans les onomatopées et dans les mots étrangers empruntés. Le *t* et le *d* sont formés en appuyant la pointe de la langue entre les dents d'en haut et les gencives et en fermant le passage avec une partie de la langue un peu aplatie. Les occlusives vélaires *k* et *g* se forment en approchant du palais la partie postérieure de la langue. Il faut remarquer que pour les occlusives brèves japonaises l'occlusion n'est pas complète et l'ouverture ne se fait pas brusquement. Les occlusives sourdes sont souvent longues, mais il n'y a pas d'occlusives sonores longues en japonais.

Les consonnes nasales japonaises ont une fermeture incomplète et un faible passage d'air par le nez. Elles ont une forte tendance à l'assimilation, par exemple *sando* « trois fois », *sammai* « trois feuilles (de papier) », *samban* « n° 3 » et *sañgai* « troisième étage ». La nasale dentale *n* est formée en japonais dans la même position que l'occlusive correspondante et la nasale bilabiale *m* se forme dans la même position que *p* et *b*. Signalons en passant les cas d'épenthèse qui se rattachent à l'assimilation ; dans la langue moderne, le mot *yūbe* « hier soir » devient *yumbe*.

Les fricatives japonaises sont des consonnes à frottement faible et elles ne présentent pas une différence marquée entre elles, surtout dans la série *h*, (*x*), *f*, et la série *s*, *ŝ*, *x'*. Les fricatives sonores *r*, *z*, *ẑ*, *y*, *w* ne sont jamais longues, non plus que les fricatives sourdes *x'*, *h*, *f*, mais *s* et *ŝ* sont très souvent longues. Dans la prononciation de *ŝ* et *ẑ* japonais on observe le phénomène caractéristique de la langue japonaise : position neutre des lèvres qui ne sont pas projetées en avant et longue fente étroite de la bouche, tandis que la langue abaisse sa pointe, et que sa face supérieure prend une position large et aplatie.

La fricative sonore *r* devant *i*, *y*, *e* est articulée près de la place d'articulation de *t* et *d*; et devant *a* et *o*, c'est un son qui est près de *l*. Les mots étrangers avec *l* sont rendus en japonais par un *r* ; la langue japonaise ne connaît pas de mots indigènes avec un *r* initial. Nous signalerons la prononciation bilabiale de la fricative *f* puis la semi-voyelle ou intervocalique *w* pour la prononciation de laquelle les lèvres ne sont pas projetées en avant et où il n'y a pas d'arrondissement et enfin une classe intermédiaire entre les fricatives et les occlusives *tš*, (*tx'*), *ts*.

L'accent du japonais est moins fort que celui de l'anglais par exemple ; l'accent d'intensité tombe souvent sur des particules *wa*, *ga*, *wo*, tandis que le mot lui-même est prononcé avec un accent monotone et une durée uniforme, mais il existe un accent musical qui est facile à percevoir dans les homophones ; par exemple : *hára* « pleine », et *hará* « abdomen », *ása* « matin » et *asá* « chanvre » ; de même des mots de trois syllabes avec l'accent sur la première, *mísete* « montrez ».

*
* *

Par son système d'écriture le japonais se rattache au chinois. Les Japonais n'ont jamais eu d'écriture originale ; dès le début du contact avec la civilisation chinoise, ils ont adopté l'écriture étrangère. C'est au commencement du ve siècle de notre ère que des savants coréens ont apporté des livres chinois au Japon et que l'étude de cette langue est devenue obligatoire pour les princes et hauts fonctionnaires. Cette écriture idéographique était très mal appropriée aux besoins de la langue japonaise. En empruntant un caractère chinois, les Japonais gardaient à ce caractère sa lecture chinoise qu'ils modifiaient d'après les lois de leur phonétique et ils lui donnaient en plus des lectures japonaises qui restauraient le sens de ce caractère : ainsi le caractère : « milieu » était lu en japonais d'après son son chinois *tšū* et en japonais *naka* « milieu », *ataru* « atteindre ».

Dans le texte le plus ancien de l'histoire du Japon « Kožiki »

compilée en 712, les scribes ont écrit le texte tantôt en chinois, tantôt en japonais en se servant pour les mots japonais de caractères chinois, indépendamment de leur signification, seulement comme d'un syllabaire phonétique. Les *Norito*, les prières rituelles, sont écrites de la même manière avec cette petite différence que les flexions des mots et les particules sont écrites en caractères moins grands. La première Anthologie des poésies japonaises, le *Manyōšū* (VIIIe siècle), est écrite en caractères chinois qui n'ont aucune valeur de signification et sont employés en tant que syllabes ; mais afin de varier l'aspect extérieur visuel pour les mêmes syllabes, on a employé divers caractères qui étaient homophones ; c'est ainsi que dans le *Manyōšū* les deux dernières syllabes du mot *tsugitete* sont rendues par deux caractères différents.

C'est seulement au IXe siècle que paraissent les deux syllabaires proprement japonais le *katakana* et le *hiragana*, tous deux dérivés d'abréviations des caractères chinois ; mais les deux écritures syllabiques représentent seulement d'une manière très approximative les syllabes réelles de la langue.

Le japonais moderne se sert d'une écriture composée de caractères chinois auxquels on souscrit au moyen de l'écriture syllabique les particules ou les éléments flexionnels. Par exemple : *To wo tatakimašita* « avait frappé la porte » sera écrit de la manière suivante : pour le *to* il y aura le caractère chinois qui représente la porte, puis en écriture syllabique sera écrite la particule du complément *wo*, puis sera écrit le caractère chinois pour « frapper » et enfin en alphabet syllabique *kimašita* pour montrer l'élément flexionnel du verbe qui est au passé avec « un élément honorifique » *masu*.

* * *

Les bibliothèques des châteaux féodaux et des monastères nous ont conservé une riche littérature. Nous avons des copies manuscrites des anciens textes des VIIIe et IXe siècles, encore peu étudiés au point de vue linguistique. Les règles de la poésie nationale et

le goût littéraire ont contribué à la formation au Japon d'une langue littéraire qui, avec certaines modifications, s'est conservée jusqu'à nos jours.

C'est vers la fin du IXe siècle que la langue parlée a commencé de s'éloigner de la langue écrite, pour en différer fortement à la fin du XIIe siècle lors des événements politiques qui mirent le centre de la vie administrative à Kamakura. Puis la langue japonaise a subi un nouveau changement lorsque les représentants de la puissance politique et militaire furent rentrés à Kyōto au XIVe siècle. Il y a eu des changements dans la langue japonaise à la fin du XVIe siècle, puis après la restauration du pouvoir impérial en 1868, lorsque quelques années plus tard Tōkyō est devenu capitale et que sa langue a servi pour la formation du japonais moderne.

Nous avons quelques textes anciens comme celui de Gokurakuganōžoka et celui de Ryōžinhisšo qui nous ont conservé la langue parlée des Xe et XIe siècles, puis le « Heikemonogatari » et le « Kanarongo » qui reflètent le parler de l'époque Kamakura (1186-1335) ; puis nous avons plusieurs textes pour la langue parlée des époques postérieures.

La langue japonaise est parlée par les Japonais habitant le Japon même : 55.961.149 habitants (1er octobre 1920), puis par quelques centaines de mille de Japonais qui habitent en Corée, en Amérique, aux îles Formose, Sakhaline, Hawaï, etc.

Serge ÉLISSÉEV.

BIBLIOGRAPHIE

Edwards (E. R.). *Etude phonétique de la langue japonaise*. Leipzig, 1903; excellent travail que nous avons beaucoup utilisé dans notre aperçu.

Balet (Cyprien). *Grammaire japonaise*, 3e éd. Tokyo, 1908 (éd. de Sansaïsha), excellent exposé de la langue japonaise parlée.

Meyer (E.A.). *Der musikalische Wortakzent im Japanischen* (*Le Monde Oriental*, Archives pour l'histoire et l'ethnographie de l'Europe Orientale et de l'Asie, vol. I, 1re partie, pp. 77-86). Uppsala, 1906.

Polivanov (F.). *Muzykal'noje, udar'enievgovore Tōkjō*. Izvestia Akademii Nauk. Petrograd, 1915.

LA LANGUE CORÉENNE

La pénurie des travaux scientifiques préliminaires, d'œuvres fondées sur la phonétique historique et sur d'autres études linguistiques, ne nous permettent pas de préciser la position du coréen dans le monde des langues. La difficulté d'une étude historique de la langue coréenne devient compréhensible lorsqu'on se rend compte que c'est seulement au XV[e] siècle que les Coréens ont inventé un alphabet qui leur a permis d'écrire leur langue.

Le coréen, par sa structure, est une des langues qu'on qualifie usuellement d'agglutinantes. Tant par son vocabulaire que par ses formes grammaticales, elle ressemble beaucoup au japonais; mais les rapports de ces deux langues ne sont pas encore précisés, et des travaux comme celui de M. S. Kanazawa : *The common origin of the Japanese and Korean languages*, Tokio, 1910, n'ont pas éclairci cette question. Les données historiques nous permettent de supposer que le coréen a subi aussi des influences du mandchou et du mongol.

Ainsi que le japonais, le coréen a été nettement influencé par la civilisation chinoise, c'est-à-dire que, à l'époque lointaine où la Chine a commencé à exercer son influence civilisatrice, la Corée était encore dans un état primitif.

Le coréen manquait d'une grande quantité de mots et c'est au vocabulaire chinois qu'il a emprunté les mots nécessaires à une civilisation déjà avancée. Il prend les noms ou les verbes chinois, et, en y ajoutant un auxiliaire *hata-* « faire », forme un verbe

coréen. Mais dans son aspect morphologique, ainsi que dans sa syntaxe, le coréen est resté homogène.

*
* *

Dans son système phonétique, le coréen ne présente pas une grande variété.

Il possède les voyelles suivantes : *a* oral libre et entravé tant long que bref *ă* ; *ə* oral libre ; *o* oral libre ou entravé ; un *u* oral libre et entravé ; un *ä* et un *e* qui sont actuellement des voyelles simples, mais qui ont gardé dans l'écriture coréenne la forme de diphtongues, et enfin un *i* oral libre et entravé (B. Karlgren, *Phonologie chinoise*, 1915, p. 299-316).

Les consonnes sont les suivantes :

p, une occlusive bilabiale orale sourde ;

m, une occlusive bilabiale nasale sonore ;

n, une occlusive dentale nasale sonore initiale et finale ;

r, qui en finale ou quand il est suivi ou précédé d'un *n* devient un *l* ;

s, qui en position finale se change en *t* ;

t, dentale dont l'articulation n'est pas exactement connue ;

tš (*č*) affriquée sourde articulée entre *t* et *k* ;

k, une plosive vélaire ;

ṅ, nasale vélaire.

Quatre consonnes *k*, *p*, *t* et *h* se trouvent aussi sous leur forme aspirée *p*', *k*', *t*' *h*', ; quatre voyelles *a*, *ə*, *o*, *u* s'emploient aussi avec un yod et deviennent *ya*, *yeu*, *yo*, *you*.

Les diphtongues de la langue coréenne présentent une certaine difficulté. La voyelle simple *ä* a, grâce à l'écriture, conservé sa forme de diphtongue. Les diphtongues *uə*, *uəi* offrent une semi-voyelle, c'est-à-dire un *u* bref ou non syllabique, qui, d'ailleurs, n'a pas été étudié dans le détail.

L'alphabet coréen comprend 11 voyelles, 14 consonnes et 12 diphtongues.

La langue coréenne possède, comme nous l'avons dit, un *a* long et un *ă* bref, qui donnent des significations différentes dans

les homophones ; d'autres homophones se distinguent d'après la place de l'accent, dont la nature n'est pas encore étudiée.

*
* *

Le coréen ne connaît ni le nombre ni le genre dans le nom commun ; le pluriel est marqué, s'il est nécessaire de le faire, par une postposition *teul*. Les noms étant invariables, c'est au moyen de postpositions que sont indiqués les rapports ; mais ces postpositions sont différentes suivant la consonne ou la voyelle qui termine la racine.

Comme postposition, on a :

pour le génitif : *eui*, *heiu*, etc. ; par exemple : « la main » *son-*, « de la main » *son-eui* ;

pour le datif, des formes : *eui-kei*, *heui-kei*, *săi-kei*, etc. ; par exemple : *son-eui-kei* « à la main » ;

pour l'instrumental : *euro*, *no*, *ro*, *să-ro*, etc. ; par exemple : « par la main » *son-euro* ;

pour l'ablatif : *ei-sye*, *kei-ro-sye*, etc. ; par exemple : « de la main » *son-eisye* ;

pour le locatif : *ei*, *xei*, *săi*, etc. ; par exemple : « dans la main » *son-ei* ;

pour l'accusatif : *răl*, *eul*, *'heul*, etc. ; par exemple : « la main » *son-eul* ;

pour le vocatif : *ye*, *ya*, *a*, *e*, etc. ; par exemple : « la main » *son-a*.

Le coréen n'a pas d'adjectifs proprement dits. Ce sont plutôt des verbes qualificatifs.

Comme adjectifs, le coréen emploie les adjectifs chinois *tai* « grand », *sio* « petit », etc., ou des adjectifs formés d'un substantif plus la terminaison *eit* ajoutée au substantif.

Les terminaisons *seurepta*, *tapta*, *haopta*, ajoutées au substantif forment des adjectifs qui ont le sens de « -able », par exemple : *sil* « vérité », *siltapta* « véritable ». D'autres terminaisons affixées

aux différentes formes du verbe donnent des adjectifs ayant le sens de « capable », « digne de », etc.

Le verbe coréen ressemble beaucoup au verbe japonais : il est impersonnel, il ne distingue pas le singulier et le pluriel ; le verbe est rarement employé avec les pronoms. Tous les rapports se distinguent par le contexte et par le caractère même du verbe qui peut être plus ou moins poli. Le verbe coréen a des affixes qui indiquent le temps, l'aspect et les autres relations des verbes ; ainsi il peut indiquer si l'on s'adresse à quelqu'un qui occupe une position sociale supérieure ou inférieure à la vôtre.

Le verbe coréen peut être divisé d'après sa conjugaison en deux classes : *a*) lorsque la conjugaison se fait par des affixes qui ont perdu leurs significations et qui sont devenus de simples parties auxiliaires et *b*) lorsque la conjugaison se fait en agglutinant des mots qui conservent leur signification propre.

En général, les formes de la conjugaison du présent, du passé et du futur sont simples et régulières ; elles finissent toujours par *a* ou *e* conformément aux lois de la phonétique coréenne. On observe la même régularité pour le participe qui est très employé en coréen.

Nous pouvons distinguer quatre conjugaisons en coréen. Les verbes avec la terminaison *ta*, par exemple : *pota* « je vois », *poatta* « j'ai vu », *pokeitta* « je verrai » ; puis les verbes se terminant en *nya*, qui est une forme interrogative, par exemple : *kananya* « vais-je ? », *kannanya* « suis-je allé ? », *kakeinnanya* « irai-je ? » ; puis la classe des verbes en *o* et en *so*, qui sont employés en s'adressant aux supérieurs et aux égaux, par exemple : *poo* « je regarde », *poaso* « j'ai regardé », *pokeisso* « je regarderai ». D'autre part il y a des suffixes pour les temps, les modes et les aspects. Les suffixes *myen* et *keteun* représentent le conditionnel, ainsi *omyen* ou *oketeun* « si je viens ». Il existe un suffixe causatif *ni*, *nikka*, par exemple : *hani* « parce que je fais cela... » Le suffixe *manan* signifie « mais », par exemple : *ponta manan* « je vois, mais... ». Le coréen a encore plusieurs autres suffixes qui donnent au verbe une signification de concession ou d'alternative « quoique » ; un conditionnel restrictif en *ya*,

par exemple : *poaya* « si je vois seulement »; puis des suffixes de temps signifiant « lorsque », « quand » et beaucoup d'autres formes compliquées.

Dans sa syntaxe, le coréen a la même structure que le japonais; le verbe est à la fin; ainsi une phrase comme : « le bœuf coréen est très grand » se traduit *ťsyosyen soka maiu k' enta* (litt. « coréen bœuf très grand est »).

La phrase a toujours un caractère descriptif, par exemple : « Il y a une belle vue du sommet de cette montagne » se traduit : *sanei olla kamyen kukyeñ mant'a* (litt. « montagne en haut si vous allez vue beaucoup »).

*
* *

Les Coréens n'avaient pas d'écriture lorsque le bouddhisme fut introduit en 372, et ce sont les caractères chinois qui ont été employés. En adoptant toute la civilisation chinoise, les Coréens ont adopté aussi l'écriture; ils ont donné seulement aux caractères un usage plus précis phonétiquement, mais ils n'ont rien changé à la syntaxe chinoise et continuent d'écrire tous leurs livres en style proprement chinois, et c'est encore en chinois qu'était rédigée l'histoire de la Corée. Au VIIe siècle, un bonze érudit, Syel-Tchong, trouva une méthode qui facilitait aux Coréens la lecture à haute voix des textes chinois en indiquant les désinences, les postpositions, les terminaisons verbales au moyen des caractères plus petits et ayant seulement une signification auxiliaire avec un phonème fixe. Ce système qu'on appelait *Nido* « la voie du fonctionnaire » ne s'est pas répandu, et on a continué d'écrire des textes chinois, sans s'occuper de la grammaire et du système coréens.

En 1443 (voir Aston, *The Onmun when invented?* dans *Transactions of the Asiatic Society of Japan*, vol. XXIII, 1895, p. 1-4), par ordre impérial, fut composé un alphabet pour noter la prononciation correcte du chinois. C'est un véritable alphabet qui a comme base l'alphabet sanskrit et où pour chaque son voyelle ou con-

sonne on emploie un caractère spécial. Les syllabes au moyen desquelles on écrit sont représentées par des caractères indépendants qui se présentent sous la forme d'une ligature composée au moyen des lettres espacées, qui se combinent en suivant l'ordre qui est la règle dans la composition des parties graphiques des caractères chinois.

« Les formes graphiques du coréen sont très faciles et logiques : les voyelles ont pour base un trait vertical ou horizontal, employé seul ou avec addition d'un ou deux traits perpendiculaires au premier et placés à droite, à gauche, au-dessus ou au-dessous; la série des labiales. . . dérive du carré; les gutturales et les dentales sont représentées par le carré privé d'un ou deux côtés. » (Maurice Courant, *Note sur les différents systèmes d'écriture employés en Corée* dans *Transactions of Asiatic Society of Japan*, vol. XXIII, 1895, p. 5).

C'est probablement en 372, avec le bouddhisme, que fut introduite en Corée l'écriture chinoise.

Dans les années postérieures, les Coréens ont écrit des notes sur les événements politiques, mais c'est beaucoup plus tard que fut écrite par ordre impérial, d'après les modèles chinois, l'histoire officielle de la Corée. Tous les livres, jusqu'au XVe siècle, ont été écrits par des Coréens, qui étaient de très bons sinologues.

Après l'invention de l'alphabet, on commença à publier des livres en caractères coréens, mais seulement pour les illettrés et pour les basses classes qui ne pouvaient pas lire les caractères chinois. Ainsi les romans, les nouvelles étaient écrits avec l'alphabet coréen, mais on ne cessa pas d'employer les caractères chinois dans la composition des ouvrages plus importants.

La langue coréenne est parlée par 16.891.289 Coréens habitant la Corée et quelques milliers de Coréens habitant la province maritime de la Sibérie.

Serge ELISSÉÈV.

BIBLIOGRAPHIE

Nous signalons la grammaire bien connue de la mission catholique française en Corée ainsi que le « dictionnaire coréen-français » contenant : I. Partie lexicographique, II. Partie grammaticale, III. Partie géographique, par les missionnaires de Corée, VIII-615-57-22 pp. et une carte de la Corée. Yokohama, 1880.

Puis le *Manuel de la langue coréenne parlée à l'usage des Français*, par G. Imbault-Huart. Paris, 1889.

La *Bibliographie coréenne, tableau littéraire de la Corée*, par Maurice Courant, Paris, 1894, donne beaucoup de renseignements.

Nous avons un bon aperçu de grammaire avec une introduction intéressante dans le livre de James Scott. *A Corean Manual or phrase book, with introductory grammar*, second edition, 1893, Seoul.

LA LANGUE AÏNOU

Les études contemporaines sur la langue aïnou ne révèlent aucune langue présentant des caractères de parenté avec elle, et s'il y a eu des langues pareilles, elles sont probablement mortes depuis longtemps ; l'aïnou est le vestige d'une langue que parlait une tribu de race blanche habitant l'Extrême-Orient. On compte actuellement 20.000 Aïnou qui parlent cette langue dans l'île Hokkadö (Yezo), dans la partie méridionale de Sakhaline (que les Japonais appellent Karafuto) et dans l'île Shikotau du groupe des îles Kouriles ; c'est aussi en raison de cette situation géographique que la langue aïnou se divise en trois dialectes, celui du Hokkadö pouvant en raison de sa phonétique être considéré comme le plus archaïque.

*
* *

Par sa structure la langue aïnou (ce nom signifie « homme ») appartient à la catégorie des langues dites de type « agglutinant », et tous ses éléments sont dépourvus de flexion.

Le nom est invariable en aïnou ; il ne distingue par la forme ni le genre, ni le nombre de l'être ou de l'objet qu'il indique ; les relations syntaxiques sont marquées par la place ou par des mots spéciaux qu'on peut appeler particules. Dans les cas où il faut préciser le singulier, on emploie l'indication numérique *šine* = « un », ainsi « un homme », *šine ainu* ; pour le pluriel on emploie les suffixes *utara*, *utari*, *utare*, — *ainu utara* « les hommes ». — *Ainu anakne ek koran* « l'homme vient » (dans cette phrase le mot *anakne* représente un nominatif emphatique) — « la maison de l'homme » *ainu kot tšisei*, — « donne au chien » *seta otta kore*.

La langue aïnou a des pronoms personnels possessifs, relatifs, indéfinis et interrogatifs et emploie des adjectifs démonstratifs. Ainsi, pour la première personne, on a en aïnou les pronoms *ku, kuani, kani, anokai,* — qui sont formés de *ku* « moi » ou « je » + *an,* qui est le verbe « être » + *i* particule transformant le verbe en substantif verbal. L'emploi du pronom est fréquent et un Aïnou dira : *kuani ku nukara* « moi je vois ».

Les adverbes de la langue aïnou sont assez nombreux ; ils peuvent être aussi formés au moyen des adjectifs en leur postposant la particule *no* : *aširi* « nouveau », *aširino* nouvellement, ou au moyen de la particule *kane* ou *koro,* qui placée après le verbe lui communique une forme adverbiale : *e* « manger », *ekane* « mangeant » ; *ahun* « entrer », *ahun koro* « entrant ».

L'adjectif de la langue aïnou est invariable et ne distingue par sa forme ni le genre, ni le nombre. Le nombre, s'il est nécessaire, peut être désigné par la particule *pa* : *pirika* « bon », *pirikapa* « bons », *pon* « petit », *ponpa* « petits ». L'aïnou a aussi des adjectifs formés par des verbes transitifs que la préposition *ši* rend intransitifs et transforme en adjectifs : *maka* « ouvrir », *šimaka* « ouvert ».

En ce qui concerne les degrés de comparaison on trouve des mots comme : « plus » *naa* et « le plus » *ijotta.* Par exemple : *pirika* « bon », *naa pirika* « meilleur », *ijotta pirika* « le meilleur ».

Le verbe aïnou peut être divisé en deux classes : d'une part les verbes qui se terminent en *ra* ou *ro* et qui changent cette syllabe en *n* devant le *n* d'un mot auxiliaire et devant *ruwe ne,* qui est aussi un mot auxiliaire donnant au verbe un sens affirmatif ; d'autre part tous les autres verbes qui ne subissent aucun changement. Ainsi le verbe *kik* « battre » reste invariable dans toute sa conjugaison ; les temps et les aspects sont formés par des mots auxiliaires : « je bats » *ku kik,* « j'ai battu » *ku kik nisa,* « je battrai » *ku kik kusu ne.*

Pour le présent, ainsi que pour le passé et le futur, le verbe aïnou a plusieurs mots auxiliaires, qui montrent les différents aspects du verbe, ainsi *ku kik širi ne* veut dire non seulement « je bats », mais « je suis en train de battre », comme en anglais « I am striking ».

Pour exprimer le passif, le verbe même et les mots auxiliaires restent invariables; seulement au lieu du pronom *ku* de la première personne nous avons : *a-en a-en kik* « je suis battu » ; à la seconde personne le verbe « être » *an* est ajouté après le verbe ou après le mot auxiliaire qui le suit : *e-kik-an* « tu es battu » et la troisième personne prend un *a* devant le verbe, *a-kik* « il est battu ».

Le verbe aïnou a des mots auxiliaires pour exprimer l'optatif, le potentiel, le concessionnel, le conditionnel et l'hypothétique.

Pour les formes négatives on trouve deux mots *šomo* ou *seenne* qui se placent devant le pronom du verbe et s'emploient dans toutes les formes : *šomo ku kik ruwe ne* « je ne bats pas ».

L'aïnou a des verbes spéciaux qui sont employés quand le complément est au pluriel : *nina* « prendre beaucoup (de choses) ».

Les noms de nombre aïnou présentent un intérêt spécial : le système décimal et le système vigésimal y coexistent. Ainsi l'aïnou compte *šine*, *tu*, *re*, *ine*, *ašikne* (*ne* est un suffixe) 1, 2, 3, 4, 5, où d'après M. Laufer le nombre 5 est apparenté au mot *aške* **ašike* « main » ; pour compter de 6 à 10 (*wan*) l'aïnou fait une soustraction 10 — 4 = 6 *iwan* ; 10 — 3 = 7 *arawan* ; 10 — 2 = 8 *tupesan* ; 10 — 1 = 9 *šinepesan* (dans les deux derniers nombres le remplacement de *wan* par un autre élément n'est pas expliqué jusqu'à présent). De « dix » à « vingt » l'aïnou compte 1 + 10 *šine ikašima wan*, etc., puis pour tous les autres nombres il emploie l'unité 20 *hotne* : 80 = 4 × 20 *ine hotne*, 100 = 5 × 20 *ašikne hotne*, 90 = 5 × 20 — 10 *wan e ašikne hotne*.

Dans la construction de la phrase aïnou, le sujet est toujours au commencement de la phrase et le verbe à la fin ; le complément est toujours avant le verbe ; toutes les propositions subordonnées précèdent le verbe principal qui est à la fin.

Le système phonique est simple. En tenant compte des sons principaux et en ne considérant pas toutes les variations dialectales, nous pouvons dresser la table suivante des consonnes aïnou :

	occlusives	nasales	spirantes
Gutturales	*k*		*h*
Palatales	*tš*(*č*)		*š*
Dentales	*t*	*n*	*s*
Labiales	*p*	*m*	*w*
Liquides	*r*		

Aisni l'aïnou n'a pas d'occlusives sonores ; les occlusives *k*, *t*, *p*, quand elles sont intervocaliques, et surtout devant *m* deviennent sonores, mais non pendant tout le temps de leur articulation. La spirante palatale *š* a tendance à devenir un *s'* palatalisé comparable à celui du polonais et sa prononciation n'exige pas que les lèvres soient projetées en avant. Comme liquide, l'aïnou a seulement le *r* qui d'après l'abbé Rousselot est une semi-occlusive ; dans quelques dialectes, *l* existe mais comme une modification de *r*.

Comme voyelles fondamentales l'aïnou a : *a*, *e*, *i*, *o*, *u*. L'abbé Rousselot indique que *a*, *e*, et *o* peuvent avoir trois qualités différentes de timbre comme en français. Laufer distingue dans le dialecte de Sakhaline huit voyelles et deux semi-voyelles.

La langue aïnou ne connaît pas d'accent d'intensité, mais il existe un accent musical qui change les voyelles brèves en longues. L'aïnou distingue *ē* « manger » et *e* « venir », *átai* « chaise » et *atái* « prix ».

*
* *

La langue aïnou n'a pas de système d'écriture, et elle ne possède pas de littérature écrite, mais le folklore est assez riche en légendes (*utšaškoma*) contes, chansons d'amour, de voyage, de pêche, etc.

Serge Élisséèv.

BIBLIOGRAPHIE

An Ainu-English Japanese Dictionary (including a Grammar of the Ainu language) by the Rev. John Batchelor. 2 vol. 1905. Tokyo.

Materials for the study of the ainu language and folklore, collected and prepared for publication by Bronislaw Pilsudski, edited under the supervision of J. Rozwadowski. Cracovie, 1912 ; excellent livre avec une bibliographie et 27 textes de contes avec des notes précieuses.

Berthold Laufer. *The vigesimal and decimal systems in the ainu numerals, with some remarks on ainu phonology*, dans *Journal of the American Oriental Society*, vol. XXXVII, p. 192-208. 1917.

LANGUES HYPERBORÉENNES

Les langues de la famille dite hyperboréenne ou « paléoasiatique » sont encore peu étudiées. La répartition en est indiquée sur la planche 6.

A l'Ouest est parlé le youkaguir (*yukagir*) ; plus à l'Est, le tchouktche (*tšuktši*) est parlé par une population de 10.000 individus qui habitent la Sibérie au Nord du fleuve Anadyr ; du fleuve Anadyr vers le Sud et le Sud-Est jusqu'à la presqu'île du Kamtchatka, se parle le koryak. Dans la presqu'île du Kamtchatka, et sur les îles Kouriles il y a une population de 2.000 habitants qui parle le kamtchadal ; enfin le guiliak est parlé dans le Nord de l'île de Sakhaline et sur le bas cours du fleuve Amour.

*
* *

Le guiliak forme les mots par juxtaposition des éléments. Le dernier mot de la série reste invariable, tandis que les mots précédents subissent des changements phonétiques, ainsi *tšaxr* « trois » et *rak* « fois » deviennent *tšrak* « trois fois ».

Le nom guiliak reste invariable ; les relations avec les autres parties du discours sont marquées par des postpositions qui indiquent les différents cas et remplacent nos prépositions : par exemple : « le bateau » *mu* ; « du bateau » (c'est-à-dire « en sortant du bateau ») *mu-x*. La postposition *-rox* correspond au datif et la postposition *-ax* à l'accusatif. Le cas peut être indiqué aussi par l'ordre des mots *itk-taf* (père-maison « la maison du père »).

Le pluriel des noms est indiqué par la postposition *xun* qui peut être omise si le sens de la phrase indique que le nom est au pluriel.

Pour les noms d'êtres inanimés, le guiliak n'a pas de genre; mais pour les noms d'êtres animés, le genre se distingue par l'addition d'un mot qui indique s'il s'agit d'un homme, d'une femme, d'un mâle ou d'une femelle, par exemple : *tšo* « poisson » *anx-tšo* « femelle du poisson » ; « l'amie » se dit : *náfx-šanx*, de *nafx* « ami » + *šanx* « femme ».

Le verbe guiliak a deux conjugaisons que M. Sternberg appelle non flexionnelle et flexionnelle.

La forme non flexionnelle conserve pour toutes les personnes et à tous les temps la terminaison de l'indicatif en *nd* et, seulement pour désigner le pluriel, prend la terminaison *xun*. Le seul changement qui soit admis dans cette conjugaison est l'intercalation d'infixes entre la racine et la terminaison *nd*. Ces infixes servent à désigner le temps (comme par exemple *i*, *in* pour le futur) ou différents aspects, etc.

La conjugaison flexionnelle a la particularité suivante : chaque temps a deux désinences, une pour la première personne du singulier et pour toutes les personnes du pluriel et une autre pour la seconde et la troisième personnes du singulier. Il existe seulement une exception, pour un temps passé, d'ailleurs peu usité, qui a une désinence pour les premières et deuxièmes personnes et une autre pour les troisièmes personnes (singulier et pluriel).

Exemple du verbe régulier de sens « marcher, aller », au parfait (temps du récit).

singulier	1re	personne	*ni vit*
	2me	—	*tši vir*
	3me	—	*hund vir*
pluriel	1re	personne	*mįzn vit*
	2me	—	*tšin vit*
	3me	—	*izn vit*

Les temps du verbe guiliak peuvent être divisés en huit classes, comprenant premièrement ceux qui ont une signification imprécise et peuvent s'employer pour rendre deux temps d'autres langues comme par exemple *ni vind* « je vais » et « je suis allé » ou *ni viind* « j'irai, je vais » ou « je vais aller » ; ensuite les temps qui ont une seule signification précise.

Comme particularité du guiliak il faut noter que le participe et l'infinitif s'accordent en personne avec le verbe qu'ils déterminent.

*
* *

Au point de vue phonétique le guiliak est assez riche.

Ses voyelles sont : *a*, *ą*, *e*, *i*, *o*, *u*, *ı*, *į*.

Le *ą* se prononce en touchant avec le bout de la langue les dents inférieures ; le son *į* se prononce comme l'*ı* profond du russe, mais avec la bouche plus ouverte en avançant la mâchoire et en abaissant la lèvre inférieure.

Les consonnes sont les suivantes : *b*, *v*, *w*, *h*, *g*, *d*, *z*, *k*, *l*, *m*, *n*, *p*, *r*, *s*, *t*, *f*, *x*, *ts*, *tš*, *š*. Parmi ces consonnes, les consonnes suivantes peuvent devenir palatales : *v*, *d*, *z*, *l*, *n*, *s*, *t*, *ts*, *tš*. On rencontre des consonnes aspirées : *b*ʻ, *g*ʻ, *k*ʻ, *n*ʻ, *p*ʻ, *t*ʻ, *tš*ʻ ; un *r* et une sorte de *g* ; une série composée des consonnes *g*, *ž*, *z*, *k*, *r*, *s*, *x*, *ts*, *tš*, *š*, avec une prononciation spéciale (ces consonnes sont prononcées en touchant les dents supérieures avec la surface du dessous de la pointe de la langue) ; des consonnes *k*, *x*, prononcées en touchant les dents inférieures avec la surface du dessus de la pointe de la langue. Ces consonnes de timbre particulier influencent la prononciation des voyelles qui les suivent, ce qui explique chez certains auteurs la notation d'un *e* particulier lequel est au fond le même *e* que nous avons noté, seulement avec un timbre différent à cause de la consonne précédente.

Le guiliak distingue des voyelles brèves et longues ; les voyelles *a*, *e*, *u*, ont souvent une terminaison assez longue qui prend le caractère d'un *h*. Les voyelles ont tendance à prendre

une intonation musicale parce que les Guiliak ont un parler chantant.

La langue guiliak n'a pas d'écriture.

S. Elisséèv.

BIBLIOGRAPHIE

Dr Wilhem Grube, *Giljakisches Wörterverzeichniss nebst grammatischen Bemerkungen* (Anhang zum III-ten Bande der Reisen und Forschungen im Amurlande von Dr Leopold v. Schrenck). Lieferung I, S. Petersburg, 1892.

L. Sternberg, *Spécimens de matériaux pour l'étude de la langue et du folk-lore des Guiliaks recueillis sur l'île Sakhaline et sur le Bas-Amour* (en russe). Bull. Acad. des Sciences de Russie. Ve série, vol. XIII, 1900, p. 387 et suivantes.

LES LANGUES PROPRES
DE
L'ASIE ANTÉRIEURE ANCIENNE

Un classement scientifique de ces langues établi soit en considération de leurs types linguistiques respectifs, soit en fonction de leurs degrés d'affinité avec des idiomes voisins ressortissant d'autres familles présenterait, cela va sans dire, un intérêt majeur.

A un tel classement, toutefois, bien des raisons s'opposent encore pour le moment :

d'abord (du moins pour une partie d'entre elles) l'insuffisance ou la pauvreté relative de notre documentation ;

ensuite — et ceci s'applique notamment aux parlers du groupe *géographique* anatolo-taurien — le caractère récent et, nécessairement, provisoire de nos informations.

L'on ne saurait, enfin, négliger de signaler la nature exceptionnellement diverse et délicate que l'étude des idiomes propres de l'Asie Mineure ancienne doit à une histoire politique passablement longue dont le détail nous échappe dans une très large mesure, mais que l'on entrevoit avoir été, par périodes, plutôt agitée, et que des facteurs économiques, des combinaisons diplomatiques, des entreprises militaires semblent avoir, comme à plaisir, compliquée.

Force nous est donc de nous en tenir au classement *purement géographique* qui présente, en tous cas, le double avantage d'être clair et de ne rien préjuger quant aux connexions éventuelles de ces parlers entre eux ; connexions soit incertaines, soit ignorées, pour le présent.

Les langues de l'Asie antérieure que l'on dénomme, par com-

modité pure, langues *asianiques* (terme qui ne préjuge aucunement de leurs corrélations mutuelles), ou *alarodiennes* (ce qui implique l'hypothèse qu'elles ont toutes pour berceau l'*Ararat*), ou *kases*, *kasiennes* (ce qui ne signifie, peut-être, pas grand'-chose), peuvent donc se répartir, à ce point de vue, en trois groupes :

A. — Basse-Mésopotamie ;

B. — Régions montagneuses périphériques et péninsule d'Asie Mineure ;

C. — Prolongements méditerranéens (parlers des îles, étrusque).

*
* *

A. — Basse-Mésopotamie.

C'est ici la région propre du sumérien.

Le SUMÉRIEN est la langue non-sémitique des plus anciennes inscriptions actuellement connues. C'est pour elle qu'a été créée, croit-on, l'écriture hiéroglyphique qui est devenue, au cours des siècles, le cunéiforme.

Cet idiome [1] était celui de tribus descendues, selon toute apparence, de l'un des hauts pays voisins, antérieurement au 5[e] millénaire av. J.-C. ; tribus qui ont donné à leur nouvel habitat sa

1. *Spécimen de sumérien* (emprunté à St. Langdon, *The sumerian Law Code compared with the Code of Hammurabi*, Journal of the royal asiatic Society, 1920, oct., p. 498) : § 3 : l. 17 à 21.

17 *tukundibi* (1) *galu* (2) si (1) [un] homme (2)
18 *giššar* (3) *galu* (4)-*ka* (5) jardin (3) d'[un] homme (4) dans (5)
19 *giš* (6) *insīg* (7) [du] bois (6) coupe (7) (ou prend)
20 *maš* (8) *mana* (9) *ku-babbar* (10) 1/2 (8) mine (9) d'argent (10)
21 *nilale* (11) il pèsera (11) (= il paiera)

(2) < à $\sqrt{gal}$ = exister ; (3) *giš* = bois + *šar* = verdure (5) ; *ka* = suffixe exprimant un sens adverbial, correspondant pour le sens à un génitif locatif ; (7) *sig* seul a le sens verbal ; *in* fait référence à un sujet éloigné = en l'espèce *galu* l. 17 ; (11) *lal* seul = *peser* ; *ni* = autre graphie de *in* ; *e* paraît être une finale emphatique.

plus ancienne organisation politique et religieuse, en même temps que les fondements de sa future civilisation.

Leur domaine, dont Nippur était le centre cultuel, et dont les villes principales étaient Kiš, Ur, Isin, Larsa, Lagaš, englobait toute la région de l'Euphrate inférieur, du site de Babylone au Golfe persique.

Nous ignorons le sens — vraisemblablement religieux — de l'ethnique *Šumer*. Nous savons seulement que les groupes de langue sémitique, successeurs politiques des Sumériens dans le pays, en ont tiré le dérivé *Šumeru*. C'est pourquoi les mots *lišan šumeri* = *langue sumérienne*, qui leur servaient à désigner ce parler, s'opposaient à celui d'*Akkadu*, qui s'appliquait à leur idiome à eux [1], que nous dénommons également *assyro-babylonien*.

L'existence d'un peuple et d'une langue *sumériens*, distincts des peuples et des langues *sémitiques* de Basse-Mésopotamie, est aujourd'hui démontrée et ne saurait plus faire doute pour personne [2].

Les monuments linguistiques sumériens consistent exclusivement en inscriptions — soit en caractères linéaires (type le plus archaïque, encore proche de l'image primitive), soit en cunéiforme proprement dit. Ces inscriptions sont gravées sur pierre — souvent très dure : diorite, serpentine, etc. ; statues, vases, cylindres, cônes, bornes, palettes ou analogues — ; ou sur pâte d'argile de forme variable : cônes, tablettes, etc. Les plus anciennes qu'on connaisse à l'heure actuelle, remontent à 4000 av. J.-C. environ.

De cette littérature, qui comprend des textes de toute espèce :

1. Cette terminologie, du fait des polémiques échangées à l'occasion de la *question sumérienne*, est restée longtemps flottante et l'on trouve des ouvrages où *accadien* = *sumérien*. L'accord, aujourd'hui, est réalisé sur la base indiquée ci-dessus.

2. Voir le Manuel de Fossey à ce sujet. — Personne ne soutient plus aujourd'hui que le sumérien ne soit qu'une autre graphie d'une langue sémitique. Mais certains assyro-babylonisants émettent encore des doutes sur la valeur des traductions actuelles au point de vue linguistique. Il ne saurait être question de discuter ici leurs arguments, mais ces réserves étaient à signaler.

économiques, juridiques, administratifs, historiques, religieux, astrologiques, mythiques, poétiques, philologiques et grammaticaux, la plus grande partie, de beaucoup (peut-être [1] même la totalité), nous a été conservée par des sémites ou, du moins, par des scribes déjà sémitisés, parlant par conséquent, et écrivant couramment aussi en dialecte sémitique.

Ceci explique pourquoi le sémitique de Basse-Mésopotamie renferme tant de traces d'influence sumérienne [2]. Inversement, le sumérien se pénètrera plus tard de sémitisme.

Langue de l'écriture, langue du culte, langue des livres, le sumérien doit à ce triple caractère de s'être maintenu pendant des siècles au milieu d'une population faisant usage d'un idiome tout autre.

A ce concours de circonstances nous devons — malgré d'inévitables fautes dans la transmission des textes — un nombre appréciable de bilingues, parmi lesquels figurent de véritables dictionnaires et des traductions juxtalinéaires.

Cette production littéraire paraît être restée relativement vivace — à peu près comme le latin au moyen âge — jusque vers 2000 av. J.-C. Passée cette date, le sumérien demeure quelque temps encore langue officielle ; mais, de plus en plus, il se confine dans son rôle de langue liturgique et savante, objet d'étude des prêtres, des scribes, des lettrés. Sous ce dernier aspect la philologie sumérienne a survécu, sans rupture de la tradition, jusqu'à l'époque des Arsacides (IIIe s. av. J.-C.).

L'on y distingue deux formes dialectales, de structure grammaticale sensiblement identique dans l'ensemble, mais comportant de nombreuses divergences de détail [3]. La plus ancienne portait le nom d'*eme-KU* ; l'autre celui d'*eme-sal*.

1. Cette forme dubitative est rendue nécessaire par notre ignorance des conditions dans lesquelles tous ces textes ont été transcrits.

2. Des traces de sumérien se retrouvent dans l'hébreu biblique également. Ex. *hēykal* « palais, temple » < sumérien : *e k/gal* « maison grande ». L'on pourra consulter à ce sujet S. LANDERSDORFER, *Sumerisches Sprachgut im Alten Testament*, Leipzig, 1916 (Beitr. zur Wiss. vom A. T. Heft 21).

3. Comparer les formes *eme-KU* : *agar* « champ » *dimer*, *dingir*, *digi-*

Ces deux formes de la langue se retrouvent dans tout son ancien domaine. Il est, dès lors, impossible d'y voir des modalités contemporaines d'un même archétype. L'on doit donc présumer qu'*eme-KU* et *eme-sal* correspondent à deux périodes nettement tranchées de son histoire. Les textes religieux, presque tous certainement postérieurs à 2500 av. J.-C., sont, pour la plupart, rédigés en *eme-sal*.

Les caractéristiques essentielles du sumérien sont :

au point de vue phonétique : la richesse relative de son système articulatoire, que la graphie cunéiforme nous dissimule, malheureusement, en partie, mais que la comparaison des variantes orthographiques permet, du moins, d'entrevoir.

En voici les principaux éléments, en l'état actuel des recherches (sous réserve des controverses pendantes) :

Voyelles ă, ŭ ĭ, (p.-ê. *ŏ* écrit *ā*), *ĕ* écrit *a-e* ou *e-a*)
Semi-voyelles : *y w*
Consonnes : *Labiales* : *p b m*
Dentales : *t d n*
Gutturales : *q k ḫ g* g^c (?) *ṅ*
Palatales : *ñ*
Sifflantes : *s z*
Chuintantes : *ś ź*
Liquides : *l r*
Liquides cacuminales : *ḷ* (?)
Liquides nasales : *l̃* (?)
Aspirées : *h* (2 variétés encore incertaines).

Il y a lieu, en outre, de remarquer la tendance marquée du sumérien à la palatalisation, à la labialisation et à l'harmonie

« brillant, dieu » ; *g'é-gal* = « excès, abondance » ; *kanam* = « pays » ; aux formes *emesal* de même sens : *adar*, *dimir* et *dimmeir*, *gemāl*, *kanaga*.

(La graphie KU indique qu'il s'agit d'un idéogramme énoncé sous sa valeur *conventionnelle*, sa prononciation *syllabique* demeurant réservée.) Le sens de ces deux termes reste incertain ; le premier, *eme*, doit se traduire « langue » ; pour *sal* et KU il y a controverse. Voir à ce propos St. Langdon, *Sum. Law Code*, pp. 3 et 4, § 2, et Haupt, *Zeitsch. für Assyr.*, XXXI, pp. 240-247.

vocalique [1] ; l'abrègement des mots par chute d'initiale [2] ou de finale [3] qui modifie parfois jusqu'à en rendre l'identification malaisée des éléments grammaticaux essentiels ;

au point de vue morphologique : la brièveté ordinaire des formes, dont un grand nombre sont monosyllabiques [4], l'absence de différenciation extérieure entre les diverses parties du discours : nom, verbe, adverbe, etc. [5]; l'absence de distinction morphologique des genres ; l'inexistence d'une flexion proprement dite ; l'emploi de préfixes, infixes et postfixes pour exprimer les rapports syntaxiques [6].

Les conséquences, *aux points de vue morphologique et syntaxique*, de la multiplicité des monosyllabes et de l'abondance extrême des homonymes se manifestent :

a) par un large recours à la composition, tant pour former des noms, que des verbes, adverbes ou prépositions [7] ;

1. C'est-à-d. à grouper dans un même mot des phonèmes vocaliques de même catégorie : soit graves = *a u* ; soit légers = *e i*.

2. Les gutturales et les nasales, notamment.

3. Même remarque.

4. De là un grand nombre d'homonymes, de sens fort différents :

Ex. : *du* = « demeure » ; « chambre » ; « bon » ; « aller » ; « apporter » ; « parler » ; « pays » ; « assembler » ; « bâtir » ; « faire cuire » ; « combat » ; « avec » ;

me = « appel » ; « appeler », « crier » ; « combat » ; « homme », « mâle » ; « ciel » ; « faire » ; « mère », « langue » ;

ku = « jeter » ; « coucher » ; « se coucher » ; « louer » (une maison) ; « arme » ; « brillant » ; « hydromel », etc. (Ces ex. sont empruntés aux lexiques de la *Gram. et Chrestom.* de S. LANGDON, et au *Glossar* de DELITZSCH).

5. En principe le vocalisme des verbes est en *e*, *i* ; celui des noms en *a*, *u*, mais les exceptions sont nombreuses ; ainsi *nad* = « être couché » ; *kar* = « tourner » ; *tar* = « couper » ; *nigin* = « chambre » ; *tig* = « cou », etc.

6. Citons, à titre d'ex. ; *r*, *š*, *da*, *ta* (les 2 premiers avec vocalisation variable soit antérieure, soit postérieure à la consonne), postfixes exprimant des rapports casuels : *mouvement vers*, *action* ou *objet*, *locatif*, *temporel*, *causal*, etc. ; *ni*, postfixe du pluriel indéterminé ; *ge* postfixe des rapports directs ; *ka* des rapports obliques ; préfixes : *al*, *an*, *at*, *am*, *mē*, etc. ; infixes : *r*(*a*), *š*(*u*), *da*, *ta*, etc.

7. Voici un ex. du procédé, d'après le *Glossar* de DELITZSCH : soit le mot *a-gu-zi-ga* = « croissance de la force » ; il se décompose en : *a* = « force » ; *gu* = « cou » ; *zig* = « lever, se lever, faire se lever » ; *a* = élément suffixé ;

b) par l'importance considérable attachée à la construction syntaxique, construction qui parvient, dans une certaine mesure, à suppléer aux insuffisances évidentes de l'appareil morphologique. — Détail à noter : le verbe se place à la fin de la phrase [1].

L'emploi de déterminatifs idéogrammatiques permet, en outre, de compenser quelques-uns des inconvénients résultant pour le lecteur de l'indétermination grammaticale et de la polysémie lexicologique signalées tout à l'heure. Sur ce point spécial, les ressemblances avec *l'usage graphique* égyptien sont parfaites.

Les obscurités fréquentes occasionnées par l'emploi d'un système d'écriture qui comporte de nombreux polyphones [2] ; les variations ou confusions d'orthographe résultant, soit des différentes écoles suivies par les scribes, soit de leur insuffisance, soit de leurs erreurs ; le rôle important des idéogrammes ; notre ignorance quant à la valeur phonétique des déterminatifs (dont on ne sait s'ils doivent ou non se prononcer) ; l'indécision qu'entraînent les *signes combinés*, dont l'interprétation prête assez souvent à l'équivoque ; l'action des tendances phonétiques mentionnées plus haut ; l'inconstance, malheureusement fréquente, de la construction ; quelque irrégularité dans l'usage des terminaisons dites *casuelles* ; l'homonymie de maints suffixes [3], donnent au sumérien une physionomie à part qui n'exclut pas une impression d'incertitude et d'obscurité.

L'extrême antiquité de cette civilisation et de ses restes suffit à rendre compte des difficultés éprouvées. Ici, comme dans les

le tout aboutissant au sens de *matin* parce que c'est le *moment où l'énergie solaire commence à croître.*

1. Le pronom personnel possessif a deux formes (indépendante et suffixale) qui sont, respectivement, pour les trois personnes sing. et plur. :

Sg.	*m(e)*	*-m(u)*	Pl. *mene*	*-mē(n)*
	zī/sī	*-zu*	*zene*	*-zune*
	nī/bī	*-ni/-bi*	*ene/bine*	*-ene* ; *-nime*, *-bime*, *-bene*

(formes du rapport *direct*).

2. Tel signe courant possède les valeurs : *g/qam*, *g/kur*, *mat/d*, *šad/t*, *lat*, *nat/d*, *k/gin*.

3. Ainsi le suffixe *-ene* du pluriel indéfini personnel est identique au suffixe possessif de sens « leur ».

Textes des pyramides, la moindre métaphore implique toute une conception du monde et de la vie à laquelle nous ne participons plus depuis bien des siècles. Cela seul constitue une grave cause d'hésitations.

Il faut tenir compte, également, de l'isolement linguistique du sumérien. Après qu'on a, tour à tour, essayé de le rattacher à l'indo-européen, à l' « ouralo-altaïque », à l'égyptien, au sémitique, au mandchou, aux langues du Caucase, ou d'y voir un idiome d'un type linguistique similaire à celui des idiomes africains, la plupart des spécialistes ont, aujourd'hui, renoncé à lui trouver des parents. La question, en réalité, n'est pas encore mûre et, sans doute, convient-il d'attendre, pour l'aborder utilement, que la plus ancienne Asie centrale et mineure nous soit mieux connue.

* * *

B. — Régions montagneuses périphériques [1] et péninsule d'Asie mineure

Mitannien. — Non loin de la Basse-Mésopotamie — où l'onomastique personnelle montre que des groupes de ce stock s'étaient introduits de fort bonne heure [2] —, se rattachant toutefois —

1. Voir le croquis à la fin du chapitre, page 309.

2. Des noms propres mitanniens apparaissent nombreux en Mésopotamie dès avant la première dynastie de Babylone (avant 2225 av. J.-C.). La présence de gloses mitanniennes dans des tablettes de correspondances écrites de Nuchassi et de Dunip en Cœlésyrie témoigne également de la présence d'un élément mitannien dans cette région. Ceci fixe approximativement les dates-limites de nos informations, soit 2200-1400 env. av. J.-C. Au sujet des noms hétéo-mitanniens en Mésopotamie voir : A. Ungnad, *Untersuchungen zu den im VII Hefte der Vorderas. Schriftdenkm., veröffentl. Urkunden aus Dilbat*, etc. Leipzig-Baltimore 1909 (Beiträge zur Assyr. u. Semit. Sprachwiss. 1909, T. vi, Heft 5.) et A. T. Clay, *Hittite-Mitannian name elements* (pp. 28-35 des *Personal Names from cuneiform inscriptions of the Cassite period*, Yale Oriental series, vol. I, New Haven-Londres-Oxford 1912) où l'on trouvera les éléments suffisants pour une bibliographie plus étendue.

bien que d'une manière encore assez mal définie — à d'autres langues de l'Asie Mineure anatolo-cappadocienne, se parlait l'idiome *mitannien* [1].

De ce parler nous ne possédons, à part de nombreux noms propres, qu'un seul texte de quelque étendue et bien conservé. Ce texte est une lettre de la collection dite de Tell-el-Amarna, lettre en graphie cunéiforme, naturellement, pourvue d'une riche vocalisation [2].

Le style de ces sortes de documents et leur plan ordinaire étant bien connus, l'on a pu, sans trop de difficultés, déterminer le sens d'un certain nombre de vocables. Tels : *ipri* = « roi » ; *šala* = « fille » ; *umini* = « pays » ; *ḫiyaroḫḫe* = « or » [3] ; *gul* = « dire » ; *ḫaš* = « percevoir ». L'on a pu, en outre, isoler certains pronoms et suffixes, ex. : *-u(š)* = 1re p. s. ; *-al* = 2e p. s. ; *-a* = 3e p. s., etc.

Comme l'on pouvait s'y attendre, le mitannien renferme des emprunts sémitiques. Ex. : *šin* = « dent » (hébreu : *šīn*, même sens) ; *šini* = « deux » (hébr. *šenāīm*). Il comportait également la présence d'éléments (sacerdotaux ?) aryens, à en juger du moins par la mention des divinités *Indra*, *Varuna*, les *Nasatyas* [4].

Le détail du fonctionnement grammatical demeure, nonobstant, assez loin d'être éclairci.

1. Le Mitanni était un petit royaume situé à l'Ouest de la Mésopotamie et qui entretenait des relations assez actives avec la confédération hétéenne, l'Arménie occidentale et la région du Liban.

2. Il est, d'ailleurs, possible que cette richesse apparente résulte en grande partie de l'écriture employée. Les syllabaires demeurés encore proches du type hiéroglyphique originel sont nécessairement de très imparfaits témoins phonétiques. Exemple : le chinois et les transcriptions grecques en caractères cypriotes.

3. C'est la forme mitannienne correspondant au grec *khrūsós* et à l'hébreu *ḫārūṣ*, termes allogènes dans l'une et l'autre langue. Noter que l'*s* intervocalique a passé à l'aspirée dans ce mot alors que celle de *Nasatya* s'est maintenue.

4. Le dernier travail sur cette question est celui de Sten Konow, *The aryan gods of the Mitanni people* (publ. de l'Institut de l'Université de Christiania), 1921.

L'une des caractéristiques essentielles de cette langue est l'emploi qu'elle fait d'éléments morphologiques mobiles qui s'accolent les uns à la suite des autres, suivant un procédé analogue à celui que l'on constate en hongrois, en turc et dans bon nombre d'idiomes du groupe caucasien.

La juxtaposition de ces *cellules* grammaticales permet d'exprimer les modalités habituelles de la pensée.

Le sens d'un bon nombre encore de ces formes demeure toutefois incertain, mais le principe paraît acquis.

Quant au système phonétique, où occlusives sourdes et sonores sont confondues, il paraît, dans son ensemble, rappeler d'assez près celui des idiomes sud-caucasiens : géorgien, mingrélien, souane, laze [1].

La grande extension des noms propres de ce groupe autorise à présumer que cette langue a joué autrefois un rôle important dans la péninsule.

Géographie et économie mercantile suffiraient, au reste, à justifier la diffusion de cet idiome, comme de l'élamite et des parlers anatolo-cappadociens. Ces peuples, en effet, témoignent par leurs archives de leur extrême activité marchande.

Au mitannien pourraient se rattacher — bien que dans une mesure impossible à déterminer pour l'instant – le cosséen et le vannique [2]. Des contaminations par voisinage demeurent, néan-

1. BORK (*Beiträge zur Sprachwiss.*, III), p. 13 et suiv., serait favorable à une parenté entre le mitannien et l'abchaze. — Il est parvenu à isoler quelques éléments, dont l'indépendance fonctionnelle paraît à peu près certaine. Citons : -*(i)kk(a)*- = élément intensif ; -*y(a)*- élément potentiel ; -*ew(a)*- = élément désidératif ; -*et(a)*- = élément inchoatif ; -*oš*- = élément qui, inséré entre le thème verbal et -*(i)kk(a)* du présent intensif, exprime le sens du prétérit (*tan-oš-ikku* : « je donnai »), etc.

2. Le seul argument que l'on puisse invoquer en ce sens est la communauté de certains éléments de noms propres. Quant aux ressemblances morphologiques, elles demeurent encore sujettes à caution et sans portée sérieuse, vu la faible proportion des comparaisons possibles. Il y a également des traits communs avec l'élamite et le hétéen. Voir WEIDNER, *Studien zur Hethitischen Sprachwissenschaft*, I, pp. 5, 10, 13, 14, 22 et suiv. Des coïncidences signalées avec ces deux dernières langues, le *parallélisme phonétique* est, de beaucoup, l'indice le plus remarquable.

moins, fort possibles, en sorte qu'il pourrait s'agir, en l'espèce, de purs et simples emprunts lexicologiques.

Du COSSÉEN [1] (on dit aussi CASSITE) nous ne possédons, d'ailleurs, qu'un maigre vocabulaire — dont l'authenticité a même été contestée par Hüsing — donnant les équivalents en sémitique d'un petit nombre d'éléments de noms propres. Tels sont, par ex. : *ašrak* = « sage » ; *bāšhu* = « dieu » ; *dakaš* = « étoile » ; *ḫamēru* et *saribu* = « pied » ; *mali* (et *meli*) = « homme, serviteur » ; *iānzi*, *nula* = « roi » ; *iaš(u)* = « pays » ; *šir* = «arc » ; *úlam* = « enfant » ; *šàribu* = « pendre » ; *uzib* = « protéger », etc.

A côté de ces termes, sans correspondants connus dans les langues soit sémitiques soit indo-européennes, l'on relève dans certains noms propres de dieux ou d'hommes des éléments formatifs manifestement aryens. Voir par ex. : *Šūmalī* (ou *Šimaliā*) = n. pr. d'une « déesse des montagnes neigeuses » (comp. le sanskrit *hima* = froid, neige et *ā-laya* = demeure, séjour = *Himā laya*) ; *šuriāš* = nom du dieu du soleil (comparer sanskrit *sūrya* = soleil).

Ces informations, par trop restreintes et limitées, en outre, au seul vocabulaire, ne nous apprennent rien sur le fonctionnement grammatical du cosséen ; langue de tribus montagnardes dont les monuments indigènes font, jusqu'ici, entièrement défaut, s'ils existent, et que l'on a tendance, en raison des éléments anaryens, mais aussi non-sémitiques, qu'elles renferment, à rattacher au groupe caucasien [2].

1. Le nom babylonien des Cosséens (grec *Kossaîoi*), est *Kaššu*. Ce peuple a prévalu en Babylonie du XVIIe au XVIe siècle env. av. J.-C. A la période cosséenne pure fait suite une période sémito-cosséenne qui prend fin vers 910 av. J.-C. L'aire d'expansion de ce peuple a embrassé depuis l'Est de la Médie et depuis l'Elam jusqu'aux environs de Diarbekir, à la Babylonie et au lac d'Ourmieh (DELITZSCH, *Die Sprache der Kossäer*, pp. 12 ; 62-63 ; 38). Les Cosséens ont conservé leur territoire propre, le Zagros, jusqu'à l'époque d'Alexandre le Grand, qui les soumit (dernier tiers du IVe siècle av. J.-C.) nominalement.

2. Voir F. N. FINCK, *Die Sprachstämme des Erdkreises* [2] (Collection *Aus Natur u. Geisteswelt*, n° 267), Leipzig-Berlin, 1915, p. 41.

Le VANNIQUE, également écrit en syllabaire cunéiforme, n'est guère mieux connu, bien que nous en possédions quelques textes un peu plus étendus.

Cette langue, qui n'offre rien de commun avec l'arménien (dont certains auteurs ont cru qu'il était une forme ancienne), nous a été conservée par un petit nombre d'inscriptions en écriture cunéiforme trouvées dans la région du lac de Van. Ces inscriptions datent des IXe et VIIIe siècles av. J.-C. et couvrent une période d'un siècle et demi environ [1].

L'emploi d'une graphie connue et la présence de nombreux idéogrammes ont permis de déterminer la teneur générale de ces textes. Mais les résultats obtenus au point de vue linguistique restent, en définitive, assez minces.

L'essentiel peut se résumer en quelques mots :

A) *Morphologie du nom* : Pas de genre morphologiquement distinct ; deux nombres (sing. et plur.) ; sept formes, différenciées par leurs finales, paraissent correspondre, *au point de vue sens*, à des cas. La variation a lieu par modification des désinences : cas sujet en *-s* ; génitif (?) en *-i* et en *-e* ; datif (?) en *-e* et en *-a* (pour les thèmes en *-a*) ; accusatif (?) ou cas régime en *-nī* ; distributif en *-nini* ; locatif (?) en *-da*, *-dai* ; une finale *-di* paraît avoir une valeur temporelle ou circonstancielle ; suffixe d'appartenance *-ni*, *-n* ; territorial *-na* ; ethnique *-a* ; d'agent *-di* ; nominatif (?) pluriel en *-s* ; cas oblique plur. en *-nie* (*nē*) et *-ni* ; cas régime pluriel en *-e*, *-ū*, *-ā*, pour les thèmes en *-i*, *-u* et *-a* respectivement ; peut-être un datif ou attributif plur. en *-ni* ; patronymiques en *-ḫini-*(*s*) ; adjectifs en *-ie*, *-ē* ; *-š* ; *-ri* ; suffixe adverbial en *-ki*(*d*) et en *-s*.

1. La langue des inscriptions vanniques est également désignée parfois sous les noms de *chaldéen* (ou *chaldique*), d'après la mention faite dans les inscriptions d'un grand dieu *Ḫaldi* ; le pays lui-même, dans la langue indigène, est dénommé *Biaina*, et, plus rarement, *Ḫaldia*. Ce pays — en totalité ou en partie seulement — est celui que les Assyriens appelaient *Urartu*, d'où le terme d'*Urartique*, usité par certains. L'Urartu paraît correspondre à l'*Ararat* des Israélites et aux *Alaródioi* des Grecs dont procède l'appellation d'*alarodienne* appliquée aux populations de la contrée : *Mosches*, *Tibaréniens*, *Táoi* (moins correctement *Táokhoi*, avec finale du plur. armén.), *Tzánoi*, *Khalubes* etc.

B) PRONOM : *mes* « lui, il » ; cas obl. *mei*; locatif (?) : *meida*, *meda* ; cas régime : *mêsini* ; relatif sg. : *ies* ; démonstratifs : *ini* (acc. sg.), *eha* « ce » ; *udas* « cela » ; *suḫe* « ces ».

C) VERBE : une forme passive paraît être formée par préfixation de *ap-* ; 1re p. sg. du passé *-bi* ; 3e p. pl. *-ni*.

D) SOUS LE RAPPORT SYNTAXIQUE, noter que le nom déterminé précède son déterminant.

Tout cela, malgré tout, reste des plus précaires.

Le vocabulaire, très réduit, ne nous apprend pas grand'chose : *ada* = « partie » ; *susis* = « brebis » ; *navū* = « chevaux » ; *manus* = « chaque » ; *kugu* = « écrire, » *zasgu* = « tuer » ; *ti* = « appeler » ; *sidis* = « de nouveau », etc.

D'après la forme des caractères, l'écriture aurait été empruntée à l'Assyrie vers l'époque de Salmanasar II (†aux environs de 830 av. J.-C.).

Faisant, en quelque sorte, pendant au vannique, mais au Sud du Zagros, dans la région correspondant au Luristan et au Khouzistan actuels, s'étendait le domaine montueux occupé par les ÉLAMITES [1].

Bien que connue depuis la découverte des inscriptions trilingues des Achéménides et désignée parfois, pour cette raison, sous le nom de *langue des inscriptions achéménides de la deuxième espèce*, cette langue n'a pu être identifiée qu'à date sensiblement plus tardive.

Aussi lui a-t-on appliqué hypothétiquement, dans l'intervalle, les noms les plus divers : *scythique*, *médo-scythique*, *mède*, *proto-médique*.

On la désigne aujourd'hui sous les dénominations *d'anzanite*, *élamite*, ou *susien*, certains préférant réserver le premier terme à

1. *Elam* = *Hatamti* en élamite ; *Nim* en sumérien ; *Elamtu* en assyro-babylonien ; *Eylâm* en hébreu ; *Elumaïs* en grec. Le mot *Anzanite* (voir ci-dessous), dérive d'*Anšan* (ou *Anzan*) qui désignait à la fois une ville et une région de l'Élam, l'une et l'autre voisines de Suse et de sa province, peut-être même identiques à celles-ci.

sa forme la plus ancienne, et les deux autres au *néo-susien*, ou *néo-élamite*[1].

Ce parler nous a été conservé par des inscriptions sur statues, dédicaces votives, documents relatifs à des constructions ou restaurations de temples, outre de nombreuses tablettes d'ordre économique[2].

Ces monuments se répartissent en trois catégories :

1° Textes en écriture indigène, dite *proto-élamite*, *élamite archaïque* ou *élamite epichorique* comportant une dizaine[3] de bilingues

1. Pour ces dénominations, voir la bibliographie donnée plus bas. M. A. H. SAYCE a appliqué à la langue des inscriptions de *Mal-Amir* — qui est un dialecte élamite assez peu aberrant, semble-t-il — le terme d'*Amardien* (< *Mardoi*, *Amardoi*) ; voir Actes du VIe Congr. intern. des Orient., 1883, Leyde, 2e part. Sect. I ; Sémit. pp. 637-756, Leyde, 1885. Certains dénomment également ce parler *médio-élamite*. — C'est A. D. MORDTMANN qui, le premier, a démontré que l'idiome des inscriptions achéménides de la 2e espèce (ainsi nommées parce que le texte figure dans la 2e colonne du texte de Behistoun) ne pouvait être que de l'élamite. Voir *Ztsch. d. d. Morg. Gesellsch.*, XV, pp. 1-126 (1862), et XXIV, pp. 1-84 (1870).

2. Notons en passant que les Élamites usaient du système décimal, à la différence des Babyloniens qui avaient adopté la méthode duodécimale. — C'est le P. SCHEIL qui, le premier, a déterminé le système de la numération.

3. *neuf*, en réalité. — Cette écriture, très proche de l'hiéroglyphe dont elle constitue une sorte d'hiératique linéaire naissant où la forme des images primitives est encore, le plus souvent, reconnaissable, paraît procéder du même archétype que l'écriture sumérienne. Elle a été délaissée de bonne heure pour le syllabaire paléo-babylonien que les Élamites ont, toutefois, adapté à leur usage. Ils ont tendu à le simplifier par l'élimination des signes *consonne* + *voyelle* + *consonne*, ne retenant que ceux *consonne* + *voyelle* (ou *inversement* (ex. : *ta-ak* au lieu de *tak*, etc.). La forme des signes a également évolué d'une manière indépendante. Aussi le syllabaire élamite est-il reconnaisable à première vue des syllabaires médio- et néo-babylonien. — Pour l'essai de déchiffrement des monuments élamites archaïques, voir C. FRANCK, *Abh. der Berl. Akad.* (Phil.-Hist. Cl.), 1912 ; *Anhang* ; *Abh.* II, pp. 1-55 et 1 pl.

D'autres documents en une autre sorte d'écriture figurative, plus ancienne, ont été découverts dans les couches profondes de Suse. Ces documents sont en voie de publication dans les *Mémoires de la Délégation en Perse* (vol. XVII). Ils ne sont pas lus, toutefois, et paraissent émaner de groupes ayant quitté l'Elam dans la suite.

(élamite et sémitique babylonien) grâce auxquels l'on a pu amorcer le déchiffrement de ce système graphique. Tous émanent d'un *patési* de Suse, *BA-ŠA-Šušinak*, contemporain, ou à peu près, de Naram-Sin (2600 env. av. J.-C.);

2° Anciennes inscriptions élamites en caractères cunéiformes élamites, dérivés eux-mêmes de l'écriture paléo-babylonienne élaborée en Élam en une graphie spéciale. Les plus anciennes datent de *Humban-umenas* et de son fils *Untaš-humban* (env. 1500 av. J.-C.). Elles se répartissent du XVIe au VIIIe s.;

3° Inscriptions élamites, également en cunéiforme, de l'époque achéménide, pour la plupart accompagnées de traductions en vieux-perse et en babylonien.

L'inscription dite de Naram-Sīn, écrite en graphie paléo-babylonienne, employée d'une manière *exclusivement phonétique* pour transcrire les sons de la langue élamite, mérite une mention à part. C'est le plus ancien texte en élamite cunéiforme paléo-babylonien.

Au cours de sa longue histoire, l'élamite a, forcément, subi un certain nombre d'altérations; altérations des plus sensibles pour l'élamite achéménide.

Celles-ci portent, comme de juste:

sur le vocabulaire, où se retrouvent des traces d'emprunts sumériens, babyloniens, iraniens;

sur la forme des mots, qui diffèrent parfois assez sensiblement, du vieux susien au néo-susien [1]. D'une manière générale, toutefois, ces modifications n'empêchent pas de reconnaître clairement les identités. Voir par ex.: v. s. *akkara* = n. s. *akkara* « quiconque »; v. s. *nap* = n. s. *nap* ou *nappi* « dieu »; v. s.

1. v. s. = vieux susien; n. s. = néo-susien. — Les emprunts au sumérien apparaissent jusqu'ici comme rares. Mais cela tient peut-être uniquement à l'antiquité de l'époque à laquelle ils ont eu lieu. L'on peut citer, p. ex. *e* = « maison »; les emprunts babyloniens et v. perse, par contre, sont relativement nombreux. Notons, en passant, que les Élamites ont appliqué à certains de ces mots le traitement grammatical de leur propre langue. Citons p. ex. le pluriel génitif *tayibuš-pe-na*, du mot v. p. *dahyauš* = pays, contrée (génitif pluriel v. p.: *dahyūnām*).

et n. s. *ú* « je » ; v. s. *aak* = n. s. *yiak* « et » ; v. s. *durna* = n. s. *turna* « connaître, savoir » ; v. s. *sunki-k* = n. s. *zunku-k* (ou *sunku-k*) = « roi » [indéterminé] ; *sunki-r* « le roi », etc.

L'appareil grammatical semble s'être assez peu transformé.

Grâce aux trilingues (vieux-perse, néo-élamite, assyro-babylonien) des Achéménides, nous pouvons nous faire une idée assez précise d'une partie du lexique néo-susien.

Ces textes, néanmoins, laissent encore subsister une certaine obscurité sur tout ce qui concerne la morphologie. Bien que de teneurs identiques dans l'ensemble, ils comportent, en effet, dans le détail d'assez nombreuses variantes d'expression d'une version à l'autre, variantes qui suffisent à faire obstacle au mot à mot intégral. L'on est donc obligé d'avoir recours également à la comparaison des variantes élamites entre elles, procédé dont les ressources sont, nécessairement, limitées.

L'examen direct et méthodique des textes de Humban-umena, Untaš-Humban, Napir-Akuš, Šutruk-Nahunte n'en a pas moins amené M. G. Hüsing à des observations importantes tant sur le développement historique de l'élamite que sur son fonctionnement grammatical.

Voici, sommairement, les traits les plus caractéristiques du néo-élamite : permutation et substitution fréquente des *sourdes* et des *sonores* : *t* et *d*, *k* et *g* ; substitution de *m* au *v* (= *w*) du vieux perse [1] ; tendance nette à la chuintisation de l'*s* (en *š*) devant occlusive; passage du *p* vieux perse à la sifflante ; chute de syllabes entières ou de voyelles à l'intérieur et à la fin des mots, probablement sous l'influence de l'accent (ce trait se trouve très accusé dans les langues du Caucase) ; pas de signe extérieur de distinction des genres ; pas de flexion proprement dite ; l'expression des rapports dits casuels (dont l'économie est très différente de celle des *cas* indo-européens) a lieu suivant deux procédés : pour le cas sujet, par sa position en tête de la proposition ; pour le génitif, le datif, le cas régime, tantôt par leur simple position

1. Citons p. ex. *Waumissa* > élam. *Maumišša* ; *Wištāspa* > élam. *Mīšdāšba* ; *Wahiazdāta* > élam. *Mīšdādda*, etc.

dans la phrase; tantôt par l'adjonction de postpositions spéciales (génitif) ou d'éléments à forme pronominale (accusatif) ; la postposition génitive est *-na* ; pour l'accusatif sg. *-in(-ir)* [1] et *-apin (-apir)* [2] ; *-itaka* a une valeur *comitative*; *-p, -pe, -ip, -ppe* = suffixe du *pluriel* des personnes.

L'ordre normal de succession de ces postfixes est : 1° postfixe de nombre, 2° postfixe casuel ; mais il y a quelques cas d'ordre inverse.

Suffixe de valeur neutre et adverbiale *-ta, -te* ; *-me* = valeur collective ; *-me, -mi* = valeur abstraite ; *-ra, -kurra*, suffixe d'agent : *(-r)ra*, ethnique ; *-umme*, suffixe numérique ordinal.

Pronoms personnels indépendants : sg. 1re p. *ú*, 2e p. *ni*, 3e p. *in, ir* ; plur. : 1re p. *nika*.

Nous avons des exemples de déclinaison (par postfixes pour les 1re et 2e p. sg.).

Possessifs : *nitami* « son » ; *nikami* « notre ».

Relatifs *akka* (plur. *akkape*) pour les personnes ; *appa*, pour personnes et choses ; démonstratifs : *ap* ; *hi* ; *hupe* ; un seul exemple de l'emploi de *ap* infixé dans un verbe comme régime direct.

Numéral : *kir* = un.

La conjugaison comporte deux types : transitif et intransitif (qui semble avoir servi à former le passif); quatre temps : une sorte d'aoriste et un parfait, un futur (?) et un présent ; cinq modes : indicatif ; impératif ; précatif ; infinitif ; participe. Sous sa forme normale le thème se présente sous l'aspect dit de l'aoriste (la finale étant vocalique).

Les désinences temporelles les mieux acquises sont : *-nta* (*-nti*) 2e p. sg. prés ou fut. ; *-nra* (*-nri*) 3e p. sg. d° ; *-niun* 1re p. pl. d° ; *-nbi, -mpi* 3e p. pl. d°.

Le présent se forme par infixation de *-ma-* entre le thème verbal et les désinences ci-dessus.

1. Les phonèmes *n* et *r* semblent avoir été voisins en élamite et permutent parfois.

2. Même remarque. L'on trouve aussi *-appin* et *-appir*.

aoriste 3 p. sg. *-š* (aussi pluriel).

-ta (*-ti*) ajouté à la forme aoristique donne le parfait : *hutta-š-ta* = il a fait ;

-ra, finale 1[re] p. sg. d'un temps préterit (?) ; 2[e] p. sg. *-š* (et plur.) ;

précatif, 3[e] p. sg. *-šne* ;

infinitif en *-mana*.

Pour la voix intransitive : aoriste 3[e] pl. sg. *-k*, *-kka*, *-ka* ; 3[e] p. *-p*, *-ppá*, *-ppi* ;

parf. 3[e] pl. *-p-ti* ;

prés. 3[e] sg. *-mak*, etc.

Cette voix est, d'ailleurs, représentée par moins d'exemples et, de ce fait, moins bien connue.

Ce qui précède peut suffire à donner une idée du fonctionnement de la langue [1]. L'économie générale du verbe, en particulier, reste, néanmoins, des plus sujettes à révision.

A ces idiomes, qui nous apparaissent jusqu'ici comme relativement indépendants les uns des autres, s'oppose le groupe des PARLERS HÉTÉO-CAPPADOCIENS ou, si l'on préfère, d'origine *anatolienne* ; parlers dont les affinités originelles, bien que non

1. *Spécimens de néo-susien* : INSCRIPTION DE L'ELWEND, DE DARIUS I (texte d'après WEISSBACH, *Die Keilinschriften der Achämeniden*, Leipzig, 1911 ; p. 100-101 ; § 1) : nāp [1] iršairra [2] uramasda [3] akka [4] hi [5] murūn [6] beišda [7] akka [8] kīkka [9] hube [10] beišda [11] ...

(un) dieu [1] grand [2] (est) Ahuramazda [3] qui [4] cette [5] terre [6] a créée [7], qui [4] (le) ciel [9] celui-là [10] a créé [11] ...

INSCRIPTION D'ARTAXERCÈS II MEMNON (Ibid., pp. 124-125, Suse b) — (cette inscription donne une idée du jeu des postfixes) :

ú [1] Irtakikšāšša [2] sunku-k [3] iršairra [4] sunku-k sunku-k-ip-inna [5] Dariiamauiš sunku-k-na [6] šā[k-ri] [7]

moi [1] Artaxerxès [2] roi [3] grand [4] roi des-rois Darius [5] — du roi [6] le fils [7].

Pour la commodité l'on a supprimé la mention des déterminatifs et groupé les éléments syllabiques en remplaçant par une longue les séquences voyelle + voyelle. L'on notera le complexe syntactique Dariiamauiš-sunkukna. Ces sortes de constructions, courantes dans certaines langues caucasiennes (ex. le géorgien), leur ont fait donner par FINCK (*Die Haupttypen des Sprachbaus*, Leipzig-Berlin, 1910), p. 144, le nom de langues à *flexion par groupes*, qui pourrait, on le voit, s'appliquer également à l'élamite.

démontrables pour le moment, paraissent, en tous cas, suffisamment probables. L'on peut donc, provisoirement du moins, les grouper sous une rubrique commune (pour le mitannien, voir p. 280) ; voir croquis plus loin, pages 299 et 309.

L'étude et la publication intégrale des nombreux textes d'ores et déjà exhumés; l'abondance documentaire que semblent promettre ces heureux débuts, permettront sans doute, d'ici quelques années, de procéder à un classement plus satisfaisant [1].

Comme l'on pouvait s'y attendre, l'idée que l'on se faisait, il y a peu de temps encore, du MONDE HÉTÉEN [2], s'est trouvée singulièrement élargie et compliquée par les trouvailles de Boghaz Keui.

Les peuples de cette région, chefs ou membres de confédérations puissantes, aux ramifications étendues, possédaient, en effet, des archives très composites qui se révèlent à nous peu à peu. D'ores et déjà, l'on y a relevé des vestiges écrits en huit langues ou dialectes différents : *sumérien* ; *assyrien classique* ; *assyrien dialectal de Cappadoce* [3] ; *hétéen* ; *proto-hétéen* [4] ; *ḫa/urri* [5] ;

1. Pour les iuscriptions en hiéroglyphes dits « hétéens », voir à la Bibliographie page 315.

2. L'on emploie fréquemment le terme anglais HITTITE, vulgarisé par les travaux de WRIGHT et d'A. H. SAYCE, pionniers de cette philologie (dans les *Proc.* et *Trans. Soc. bibl. Arch.*).

3. Cette langue, comme on le sait, n'est qu'une forme dialectale de l'assyrien où l'influence cappadocienne est particulièrement sensible. Pour la bibliographie de ces textes d'ordre tout particulier, voir JOHNS, *Schweich Lectures*, Londres, 1912, pp. 88 et suiv. G. CONTENAU, *Trente tablettes cappadociennes*, Paris, 1919 ; les publications de tablettes cappadociennes les plus récentes sont celles de G. CONTENAU, *Tablettes cappadociennes publiées avec inventaire et tables* (Musée du Louvre, Dépt. des Antiq. orient., t. IV), Paris, 1920, des *Cuneiform texts from Cappadocian tablets* (Part I), British Museum, Londres, 1921, et de E. F. WEIDNER, *Die Assyriologie*, 1914-1922 (Lpz., 1922), pp. 187-8.

4. Ces noms et ceux qui suivent, noms tout provisoires, sont ceux de l'opuscule de M. FORRER. Il dénomme *proto-hétéen* une langue qu'il estime avoir été originellement celle de la majeure partie du pays hétéen. Quelques bilingues lui ont, parait-il, permis d'en interpréter des textes. Le type de ce parler — à préfixes et postfixes — paraîtrait voisin du géorgien. La grammaire en serait relativement compliquée et du modèle caucasien.

5. Suivant WINCKLER, *Or. Lit. Ztg*, 1910, p. 281 et suiv., ce nom serait

lūi [1] ; *palawi* [2], *manda*, enfin, où E. Forrer penche à voir un nouveau dialecte indo-européen ; et rien ne prouve que cette liste doive désormais rester close [3].

Les monuments en hétéen comprennent, à ce jour : des vocabulaires (suméro-assyro-hétéens) ; des instruments diplomatiques ; un code ; un traité technique relatif au dressage du cheval ; un poème théologique, variante, semble-t-il, de l'épisode babylonien de Gilgameš ; des fragments d'incantations et de liturgies, plus des pièces d'importance et d'intérêt variables : contrats, lettres, etc., le tout se chiffrant par plusieurs milliers de tablettes.

La graphie cunéiforme dont il est fait usage [4] pour tous ces textes a permis de les transcrire — du moins en partie —, l'expression phonétique des idéogrammes demeurant, comme de juste, réservée.

l'ancienne forme de celui d'*Aryens*. M. A. H. Sayce, *Journ. of the roy. As. Soc.*, 1920, pp. 49-83, préconise la lecture *murri* ; l'on a retrouvé en cette langue des portions du poème dû à un certain *Kesse*, comportant 14 tablettes, chacune à deux colonnes de texte, relatives à l'histoire d'un dieu *Kumervi* (?) et de *Galgamis* (probablement le *Gilgameš* babylonien).

1. Lu aussi *luwi*. Sayce propose d'y voir un hétéen littéraire, plus pur (art. cité à la note précédente, p. 65). Forrer distingue nettement cette langue du hétéen de Boghaz-Keui.

2. En cette langue nous ne possédons jusqu'ici que des incantations. Selon Forrer elle aurait été parlée à *Palâmu*, ville qu'il situe au Nord-Est de l'Asie Mineure, entre Tokat et Sivas, au Sud du Yildiz Dagh.

3. Au nom de *Hétéen* le même Forrer propose de substituer le nom de *Kanèsien* se fondant sur ce qu'il est fait, dans certain fragment de liturgie, mention du chanteur de *Kaneš* outre le chanteur de Ḫattu. La langue désignée jusqu'ici sous le nom de hétéen, et qui est le plus largement représentée dans les monuments découverts, ne serait pas le hétéen. Aussi a-t-on proposé de l'appeler pseudo-hétéen (pseudo-hittite), ce qui n'est pas très clair (voir encore ci-dessus, p. 21). — Pour apprécier la valeur et la portée linguistique de tous ces termes, il y a lieu d'attendre une publication *intégrale* des textes découverts.

4. Le choix des signes et des valeurs syllabiques qui leur sont attribuées, ainsi que l'allure générale de la graphie permettent de fixer vers l'époque des dynasties d'Isin (aussi *Nisin*) et première dynastie de Babylone (soit, approximativement, entre 2350 et 2200 av. J.-C.) la date à laquelle l'écriture cunéiforme hétéenne s'est constituée en système distinct.

Ces derniers signes, passablement nombreux, ont néanmoins fourni quelques indications intéressantes sur la teneur générale de ces matériaux.

Au point de vue strictement linguistique, cependant, le seul qui nous occupe ici, le résultat obtenu, demeure, malgré tout, relativement précaire.

Des mots qu'on lit sans les comprendre, pour la plupart ; des signes que l'on comprend, mais sans les lire [1] ou en ne lisant, en bien des cas, que leurs *compléments phonétiques*, constituent une ressource assez modeste.

Cette considération seule semblerait devoir suffire pour inciter à quelque réserve.

Pour le moment, nos sources d'informations précises peuvent, en somme, se grouper sous deux chefs principaux :

A : les vocabulaires ;

B : les textes suivis.

A. — Les premiers nous apprennent, il est vrai, peu de chose sur le fonctionnement même de la langue puisqu'ils ne nous fournissent, par définition, que des vocables isolés.

Ils nous apprennent, p. ex. : que l'hétéen *i-ya-o-wa-ar* = assyr. *e-pe-šu* « faire » ; que *pe-tar* = *ka-ppu* « aile » ; que *iš-har* = *da-am-mu* « sang » ; que *i-ta-lu-uš* = *pa-aš-qu* « mauvais », pénible ; que *ta-at-ta-lu-uš-ki-o-wa-ar* = *pe-du-ú* « libérer, délivrer, acquitter », etc., etc.

Ces indications, si modiques soient-elles, sont précieuses. Elles nous fournissent, en effet, des certitudes immédiates quant au sens et à la forme d'un certain nombre de mots. Elles nous soustraient, en ce qui concerne lesdits mots, à toutes spéculations étymologiques.

Quelle que soit l'opinion que l'on professe ou que l'on préfère touchant la véritable nature de « l'hétéen », les faits ci-dessus demeurent acquis.

Les vocabulaires nous valent, par surcroît, une donnée

1. V. A. MEILLET, *Bul. de la Soc. de Ling. de Paris*, 1920, t. XXII, fasc. 1, p 112.

phonétique intéressante : l'absence en « hétéen » de véritables sonores. Ce caractère, qui se remarque en sumérien, en mitannien, en élamite, se retrouve dans les langues du Caucase. Il se rencontre aussi en étrusque, que l'on tend aujourd'hui à rattacher au groupe caucasien [1], et que la tradition, en tout cas, relie formellement à l'Asie Mineure.

Certains des mots déchiffrés, assurément, paraissent pouvoir se rapporter à des éléments ou à des racines indo-européennes de sens analogue [2]. *A priori*, la présence de contingents aryens en Asie Mineure — ou celle de divinités aryennes en pays mitannien — n'a rien qui puisse nous surprendre aux environs du XVe s. av. J.-C. Le contraire seul aurait lieu de nous étonner. Dans son ensemble, néanmoins, le lexique paraît comporter d'importants éléments étrangers aux langues indo-européennes connues.

Les vocables précités n'autorisent donc, *à eux seuls*, aucune conclusion ferme quant au fond des choses. Les mots sémitiques en turc, vieux-perses en élamite, sanskrits en chinois n'empêchent pas le turc, l'élamite, le chinois d'appartenir à des systèmes distincts des systèmes sémitique et indo-européen.

Quelques analogies que l'on a constatées avec le latin, p. ex. entre *uerite* = craindre, peur, et *uer-eor* = je crains, ne sont guère plus démonstratives *au point de vue strictement indo-européen*.

Pas davantage celles de *mi* « je, mien » car le géorgien, qui n'est pas indo-européen, dit *me* pour « je, moi ».

Une véritable certitude ne saurait résulter que d'une plus intime connaissance de la grammaire « hétéenne ». C'est ici qu'intervient, naturellement, le secours des textes suivis.

1. V. par ex. F. N. FINCK, *Die Sprachstämme des Erdkreises*, Leipzig-Berlin, 1915, p. 41.

2. Comparer l'exemple de l'hétéen *war* = assyr. babyl. *mu-u* = eau, que WEIDNER, *op. cit.*, rapproche, non sans vraisemblance, de l'avare *hor* et du basque *ur* (même sens), mais que l'on peut également conférer, avec le sanskrit *vāri* = eau, tokharien *wär*. Ceci suffit, pensons-nous, à illustrer les risques de comparaisons portant *exclusivement* sur le vocabulaire.

B. — Nous avons signalé plus haut les raisons qui rendent difficile d'utiliser ces textes.

Pour le moment, on peut s'en tenir aux constatations suivantes — constatations, du reste, toutes provisoires, en raison du grand nombre de textes restant à publier et du caractère *éminemment international* de la confédération hétéenne — :

Vocabulaire. Présence d'éléments sémitiques[1], d'éléments indo-européens (aryens)[2], probablement d'éléments paléo-caucasiens (?)[3] ou que leur nature autorise, du moins, à suspecter d'affinités originelles avec le monde caucasien (mitannien et élamite).

Grammaire. Voici ce que permettent d'entrevoir les vocabulaires précités complétés par les indications du Code récemment publié par Hrozný. Ce dernier document forme, en effet, un tout homogène, relativement peu chargé en idéogrammes, où le parallélisme des formules nous vaut une appréciable sécurité. Les sens sont, bien entendu, sujets à révision là où ils ne sont point garantis par les vocabulaires.

Nom, singulier :

Cas sujet : en *-a*, *-i*, *-u* ; *-aš*, *-iš*, *-uš* ; *-n*, *-t* ; *-ar* ; *-l*
Cas régime : en *-an*, *-in* ; peut-être en *-uš* ;
Datif (?) : en *-i* ou *-ni* ;
Ablatif et/ou instrumental en *-az* ;
Génitif : *-aš* (fém. sing.) ;
Locatif (?) : en *-i*.

pluriel :

Cas sujet : *-iš* ; *-eš* ;
Cas régime : *-an* ; *uš* ; *-a* ;
Cas oblique (?) *-ás*.

Pronom.

ku-iš, *ku-it* « qui, que » (cas sujet) ;

1. P. ex. l'assyro-babylonien : *ul* « ne... pas ».
2. C'est-à-dire indo-iraniens.
3. Weidner, *Studien zur hethit. Sprachwiss.*, I, p. 34-35.

ku-iš-ki, ku-it-ki, ku-iš-ku-iš, ku-it-ku-it « quelqu'un ; quelque chose » (cas sujet) ;

a-pu-un « celui-là » (cas régime) ;

ta-an « celui-là » ou « -ci » (cas régime) ;

mi « je, moi » (cas sujet) ;

-mi, -mi-iš « mien » (cas sujet) ;

ši = lui (cas sujet) ; *-ši* « son » (cas sujet) ;

naš, nat « lui, cela » (cas sujet) ;

šu-ra-aš « à vous », ou « votre ».

Verbe.

Présent de l'Indicatif (ou équivalent) :

3ᵉ pers. sg. en *-i* ou *-zi* ;

3ᵉ pers. plur. en *-án-zi* ;

Prétérit (?) :

3ᵉ pers. sg. *-š-ki-ir* : Ex. : *pa-a-i* « il donne » ; *pi-eš-kir* « il donnait » ; *da-a-i* « il prend » ; *da-aš-ki-ir* « il prenait » ;

Aoriste (?) :

3ᵉ pers. sg. *-it* : *pi-eš-ši-it* « il a donné » ;

Participe passé adjectival :

singulier : *-an-za* : *ḫu-u-ni-in-kán-za* « le frappé (?) » (cas sujet) ;

pluriel : *-a-an-te-eš* (= *-anteš*) : *ta-ru-up-pa-an-te-eš* « les réunis » (cas sujet) ;

Infinitif en *-ar*.

Adverbes et/ou prépositions ou analogues :

ták-ku « si » ; *na-aš, na-an* « et » ;

na-aš-ma « ou bien » ; *nu* « or donc, maintenant » ; *-na-ku* « ou bien », enclitique du type du *-ve* latin ; *-pe* sorte de deictique analogue au *-pse* latin dans *i-pse, ea-pse*.

On le voit : si le vocabulaire nous reste en grande partie obscur, la morphologie est de physionomie indo-européenne marquée.

Il convient, toutefois, de garder présente à l'esprit l'existence entre le hétéen et les idiomes caucasiens — (dont la conformité générale de structure avec le mitannien et l'élamite constitue

un indice remarquable) — de coïncidences, notamment *phonétiques*, dont la portée reste à déterminer.

La mêlée des peuples que Strabon [1] signale en Asie Mineure après la guerre de Troie paraît y avoir été un phénomène fort ancien. Il faut donc se garder de tout système anticipé et savoir attendre patiemment l'épreuve des faits.

Au hétéen se rattachent, par des liens qu'il est encore difficile de déterminer scientifiquement, mais dont la réalité parait difficilement contestable [2] :

les IDIOMES ASIANIQUES CÔTIERS DE L'ASIE MINEURE : *cilicien*, *carien*, *lycien*, *pisidien*, *isaurien*, *lycaonien*, *cataonien*, *lydien*, *mysien*, *gergithe-solyme*.

La plupart de ces langues ne sont plus pour nous que des noms et nous ignorons jusqu'à quel point ces appellations correspondaient, en ce qui concerne le plus grand nombre d'entre elles, à une différenciation linguistique effective.

Du cilicien, pisidien, isaurien, lycaonien, cataonien, gergithe-solyme, il ne reste rien ou, tout au plus, que deux ou trois gloses isolées dont il y a peu à tirer.

Certaines tendances articulatoires ont, néanmoins, subsisté sur les anciens terroirs de ces langues, tendances dont l'étude systématique à travers le grec des inscriptions comporte nombre d'utiles indications.

Le CARIEN, un peu moins mal connu, est représenté par d'assez nombreux noms propres du type spécial qui se retrouve dans toutes les régions adjacentes; par quelques vocables, attestés presque tous chez Étienne de Byzance [3] ; enfin, par environ quatre-vingts inscriptions ou graffiti, tous fort courts, dont trois bilingues égypto-cariens.

1. STRABON, XIV, C. 678.

2. L'onomastique y est, en effet, d'un type bien spécial et aisément reconnaissable. Ceci implique en tous cas une aristocratie commune, ce que confirment les alliances de famille. Ainsi Pixodaros de Carie épouse la Cappadocienne Aphnéis (STRABON, XIV, c. 657).

3. *Ala* = cheval, *banda* = victoire, *soua* = tombeau, *gela* = roi, etc.

Ces monuments sont écrits en caractères alphabétiques analogues aux lettres grecques et se rattachent, sans aucun doute, à un archétype graphique commun ; mais les valeurs d'un certain nombre de signes propres au carien ne sont pas encore établies d'une manière sûre [1].

Le lydien et le lycien sont mieux attestés, bien que par des inscriptions de date également récente, sans compter des gloses transmises par les auteurs et lexicographes classiques.

Les principaux textes LYDIENS sont dus aux fouilles de l'*American Society for the excavation of Sardis*. Un certain nombre d'entre eux, dont un bilingue araméo-lydien, malheureusement assez peu clair dans sa partie araméenne, ont été publiés. D'autres le seront prochainement.

L'on a pu établir la valeur de la plupart des lettres — pour la plupart identiques aux lettres grecques. Les noms propres ont permis de déterminer la valeur des signes spéciaux au lydien, dont trois ou quatre, toutefois, n'ont reçu qu'une expression phonétique provisoire.

Ce système articulatoire se distingue du grec par l'absence d'occlusives aspirées (χ, θ, φ) et par celle de l'occlusive labiale (*p*) [2]. Il comporte, en outre, comme le lycien, un certain nombre de phonèmes nasalisés, de type soit *vocalique* soit *sonant*, dont la véritable nature n'est pas encore bien fixée.

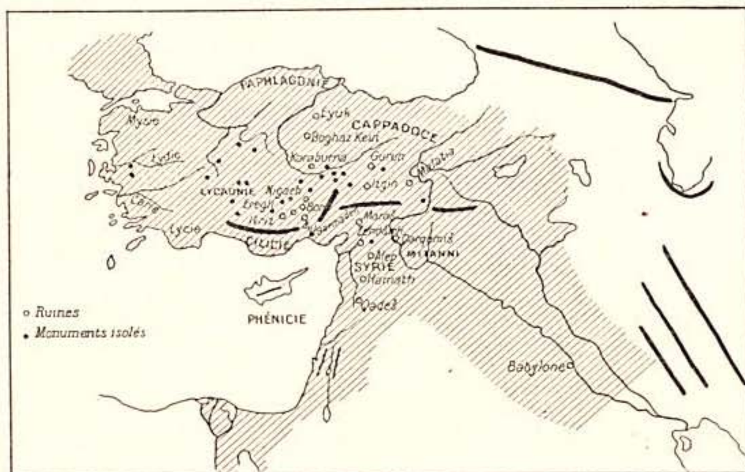
Au point de vue morphologique, on remarque l'absence de signe extérieur de distinction des genres ; l'usage de *-s* et *-l-*

1. La date des inscriptions cariennes est relativement récente ; les plus anciennes sont contemporaines de Psammétique, soit 667 env. av. J.-C. Comme tous les peuples de la mer, les Cariens se sont diffusés sur une grande étendue, du Pont à la Grèce et à l'Afrique du Nord (> Gibraltar).

2. Voici un spécimen de lydien, suivant la transcription d'E. LITTMANN, *Sardis*, p. 54, accompagnée de la traduction dudit. (Les numéros identiques indiquent le mot à mot suivi :) ll. 1-4 : *eśś* (1) *vã(n)aś* (2) *Sivam̃lis* (3) *Armãvlis* (4) *akit bis esũ* (5) *vãnaũ* (6) *buk* (7) *esćać* (8) *antolać* (9) *buk* (10) *esćać* (11) *labrisać* (12) *fènsũibid* (13), etc. — Cette (1) tombe (2) [est] de Sivãm (3) fils d'Ar-

comme éléments casuels et de dérivation ; la présence d'une conjonction enclitique *-c*, analogue à l'enclitique étrusque *-c*. Les

Croquis indiquant les localités où la présence de monuments dits « Hétéens » permet de présumer l'usage, à tout le moins comme langue officielle, de l'un des idiomes usités dans la confédération cappadocienne pour la rédaction des actes publics [1].



Les hachures correspondent à la diffusion approximative des noms asianiques en Asie Mineure.

mãv (4) si quelqu'un cette (5) tombe (6) ou (7) ces (8) corps (?) (9) ou (11) lits funèbres (12) (?) détruit (13), etc.).

Les voyelles surmontées du tilde sont nasalisées : la valeur du *ś* et du *ć* reste douteuse.

Les inscriptions lydiennes actuellement connues datent des Ve-IVe s. av. J.-C.

A la place du *p*, le lydien comporte une spirante très douce notée *f*, qui paraît s'être approchée plutôt du *w* que du *f* proprement dit. C'était, en tous cas, une labiale fluante.

1. D'après : A. T. Clay, Personal names of the Cassite period, 1912. — A. Deimel, Pantheon Babylonicum, 1914. — P. E. Huber, Die Personennamen in den Keilschr. Urkunden aus der Zeit der Könige von Ur und Nisin, 1907. — J. Lieblien, Hieroglyphische Namen, dans Wörterbuch, 1910. — H. Ranke, Early Babyl. Pers. Names, 1905. Continué par Hilprecht. — W. Budge, Egypt. Dictonary, 1920. — Tallqvist, Neubab. Namenbuch. — Luckenbill, Am. Journ. Sem. Lang., XXVI. — Sundwall, Klio, 1913.

monuments publiés ne permettent pas d'avoir une idée claire du fonctionnement du verbe.

C'est tout ce que l'on peut dire de certain pour l'instant.

Du LYCIEN les vestiges épigraphiques, également écrits en caractères alphabétiques, en partie identiques aux caractères grecs, en partie différents, sont relativement nombreux. Quelques-uns, telles les inscriptions de Xanthos, d'Isindis, de Limyra, sont même assez étendus. Mais les bilingues helléno-lyciens sont courts ou mutilés, donc assez peu propices à la reconstitution de la grammaire[1].

Ils ont permis, néanmoins, de reconnaître avec certitude le sens d'un certain nombre de mots, tels : *tideimi* = fils ; *lada* = femme, épouse ; *prñnava* = tombe, demeure ; *kupa* = caveau, fosse ; *ebẽñne* = pronom démonstratif ; *ehbi* = pronom possessif de la 3ᵉ pers. sing. ; *ti* = relatif ; *tike* = pronom indéfini ; *hrppi* = préposition correspondant pour le sens à grec *epi* ; *se* = et, etc.

Le système phonétique du lycien comportait, du moins autant que l'écriture le laisse discerner, un minimum de 29 phonèmes distincts.

Ce système n'était que très imparfaitement superposable au système articulatoire grec. Cette inadaptabilité réciproque tient en grande partie au caractère instable (au point de vue hellénique) de certaines voyelles lyciennes (notamment de l'*e*) ainsi qu'à la présence en lycien de phonèmes *sonants* et *nasalisés* auxquels nous avons fait allusion tout à l'heure.

Les flottements qui en sont résultés dans les transcriptions hel-

1. Les inscriptions lyciennes ne remontent pas au delà du Ier s. av. J.-C. — Voici un spécimen de la langue : *Tituli Asiae minoris* 6 (bilingue gréco-lycien) : *ebẽñnẽ* (1) *ñtatã* (2) *me* (3) *ne* (4) *prñnawatẽ* (5) *pulenjda* (6) *mullijeseh* (7) *se* (8) *dapara* (9) etc.

ce (1) monument (2) ci (3) [est] du (4) maître-de-maison (5) Apollonidès (6) [fils] de Mollisis (7) et (8) Laparas (9), etc. (Pour le mot à mot v. AUTRAN, *Tarkondemos*, pp. 129-139).

Les caractères surmontés du tilde représentent des sons nasalisés.

léniques nous valent, d'ailleurs, bon nombre d'indications intéressantes sur les procédés propres du lycien.

Les occlusives, notamment les labiales, étaient certainement assez faibles, particularité qui, en Asie Mineure, ne saurait guère surprendre. Il semble que cet idiome témoigne également d'une tendance vers l'harmonie vocalique.

Grâce, surtout, aux noms propres, il a été possible de reconstituer les principales modifications désinentielles du substantif correspondant à la déclinaison.

L'on a reconnu l'existence de thèmes en *-a*, en *-i*, en *-u* et en *-s*(*h*); ces finales sont celles du nominatif singulier. Le génitif est en *-ah*(*e*), *-ih*(*e*), *-uh* pour les trois premiers ; le datif en *-i* et *-aye*, et *-i* ou *-eye* pour les deux premiers; l'accusatif en *-ã*, *-u* ou *-a* pour le premier, *-i* ou *-iyẽ* (douteux) pour le deuxième; *-sñ*(*hñ*) pour le quatrième. Les thèmes en *-a* forment leur pluriel nominatif en *-ahi* (douteux), datif en *-a* ; accusatif en *-as* (ou *az*); ceux en *-i* en *-iye*(*a*)*hi* (?) ou *-i* ; *-e* ou *-iye*; *-is* (ou *-iz*) ou *-iye*(*a*)*his* (douteux) respectivement.

Le fonctionnement du verbe, par contre, nous échappe à peu près complètement. L'on peut, cependant, mentionner une désinence verbale *-ti* (3[e] pers. sing.) qui revient assez fréquemment dans les textes [1].

*
* *

C. — Prolongements méditerranéens (parlers des iles; étrusque)

Nous en avons terminé avec les langues propres de l'Asie Mineure ancienne, du moins avec celles parlées dans la péninsule même. Pour être complets, il ne nous reste plus que quelques mots à dire des survivances et prolongements méditerranéens desdits parlers.

Comme on le sait, ces survivances sont, pour la plupart, très

1. *Prñnawatẽ* n'est pas un verbe. Voir *Tarkondemos*, ch. III, pp. 130 et suiv.

fragmentaires, mais la seule onomastique suffit à démontrer que l'influence asianique était prépondérante dans ces idiomes. Le témoignage des lexicographes est également en ce sens. La tradition historique n'est pas moins formelle.

De ces vestiges linguistiques, le plus important de beaucoup, tant à raison du nombre que de l'accès facile et du bon état de conservation de ses monuments est, sans contredit :

L'ÉTRUSQUE.

Ce n'est point ici le lieu d'exposer les raisons de tout ordre : archéologique, historique, économique, religieux [1], linguistique qui autorisent à considérer comme certaine une affinité entre la langue de la dodécapole étrusque et celles de la côte d'Asie.

Les difficultés commencent lorsque l'on veut chercher à déterminer la nature et la portée précise de cette affinité, car nos informations sont encore assez pauvres.

De plus, les conditions politiques prévalant en Méditerranée lors de la venue des contingents étrusques en Italie ainsi que la date présumable (deuxième moitié du 2ᵉ millénaire av. J.-C.) de cette arrivée incitent à penser que l'étrusque ne pouvait pas ne pas être une langue des plus adultérées.

Le peu que l'on en sait peut actuellement se résumer en ceci :

au point de vue phonétique : absence de vraies sonores et type spécial d'occlusives à occlusion très faible que les alphabets grecs et italiques rendent indifféremment soit par des douces, soit par des fortes, soit par des aspirées, des spirantes et — pour les labiales — par la semi-voyelle du même ordre ; existence possible d'une sorte d'apophonie vocalique [2] ; tendance [due à l'accent (?)] à l'abréviation des formes [3] et à l'accumulation des consonnes ; large usage de la prothèse et de l'épenthèse ; tendance à la

1. V. la bibliographie étrusque et lydienne.

2. Comparer p. ex. *clan* = fils ; pluriel *clen-ar*. L'on trouve, toutefois, la forme du sing. *clen* (CIE 446 Cortone) et le génitif *clenś* (CIE 4050 Pérouse).

3. Citons p. ex. des abréviations telles que *Achille* > *Aχle* ; *Clytemnestre* > *Clutmsta* ; *Alexandre* > *Elaχsantre* > *Elsntre*, etc. La question de l'accent mérite toutefois quelques réserves.

nasalisation [1] ; à la diphtongaison [2] ; obscurcissement des voyelles.

sous le rapport morphologique : emploi de morphèmes et ou de suffixes, dont la valeur n'a pu, dans la majorité des cas, être exactement fixée.

L'on constate, néanmoins, des gentilices en *-tru*, *-tra*, *-ma*, *-ra*, *-la*, *-sa*, *-tla* ; etc. ; en *-es*, *-us*, *-o*, *-iu*, *-sa*, *-θur* ; la finale *-na* (et *-enna* [3]) y fait même l'objet d'une prédilection toute particulière ; un suffixe *-aχ* caractérise les ethniques [4].

Certaines désinences spéciales servent à exprimer les rapports dits casuels ; ainsi *-e* se trouve au nominatif singulier d'une catégorie de noms [5] ; *-er* pour un autre groupe ; *-ni* pour un troisième ; *-s* (ou *ś*) [6] pour un quatrième [noms masculins] ; pour le féminin, l'on trouve *-a*, *-i*, *-ia*, *-an*, *-θa* (*-ta*) [7] ; *χ* (et *-χi*) a une valeur multiplicative ; *-(a)r*, *-(e)r* est la désinence du pluriel et entraîne en certains cas une apophonie de la voyelle du thème ; le génitif singulier est en *-s* (ou *-ś*) ; le datif (?) sing. en *-śi* ;

1. Voir les remarques du CIE, I, p. 59 avec les exemples.

2. Citons des variantes comme *Sature*, *Satrna*, *Satrenus* et *Sautarine*, *Sauturinial*, *Sautri* ; *Rufi*, *Rafi* et *Raufi*, *Raufe* ; *Claniu* et *Clauniu*, etc. L'on trouvera bien d'autres exemples dans W. SCHULZE, *Zur Gesch. lat. Eigenn.* Abh. Götting., 1904.

3. Cp. *Porsenna*, *Ravenna*, etc., et cp. en Asie Mineure, *Etenna*, *Trebenna*, etc.

4. Citons par ex. : *rum-aχ* = Rom-ain. Il existe également un autre suffixe ethnique *-al* qui paraît se retrouver avec la même valeur dans l'idiome de l'inscription de Lemnos.

5. Citons par ex. : Tydée, grec *Tudeús* = étr. *Tute*. Cette désinence paraît avoir été caractéristique de l'étrusque.

6. Ces deux phonèmes représentent deux variétés mal définies de sifflantes dont l'une prévaut dans l'Étrurie du Nord, l'autre dans celle du Sud, mais qui permutent constamment dans les inscriptions.

7. Il est possible que ces désinences féminines ne soient pas originelles en étrusque et soient dues à l'influence des dialectes italiques. Nos plus anciennes inscriptions étrusques ne paraissent pas pouvoir remonter au delà de 600 av. J.-C. L'étrusque était encore pratiqué par les haruspices au temps de Julien l'Apostat (Ammien XXIII, 5) = 360 ap. J.-C., donc réduit à la fonction d'idiome liturgique. Quant aux inscriptions étrusques, elles s'arrêtent vers le siècle d'Auguste.

un génitif-ablatif sing. en *-al* ; un génitif (?) anomal non expliqué en *-um* ; à relativement basse date, le génitif en *-s* (*-ś*) s'accole un suffixe supplémentaire (probablement analogique) en *-la* (ou *-iśla* après un génitif-ablatif en *-al*), d'où une forme nouvelle désignée d'ordinaire par le terme *genitivus genitivi*, d'emploi très fréquent pour les patronymiques.

Les principaux démonstratifs sont *mi* et *cen*.

Les noms de nombre sont *maχ* = 1 ; *zal* = 2 ; *θu* = 3 ; *huθ* = 4 ; *ci* = 5 ; *śa* = 6 ; *-alχ* exprime les dizaines. Ex. : *ce-alχ* = cinquante.

Le verbe, très mal entrevu, comportait, entre autres, une désinence *-ce* comme indice de la 3e pers. sing. d'un temps passé. Cette désinence, que l'on retrouve accolée à des titres de fonctions officielles exercées par des défunts [1], prouve l'existence en étrusque de formes verbales dénominatives du modèle de nos : *solutionner*, *occasionner*, etc.

L'on a pu isoler deux conjonctions enclitiques *-c* [2] et *-m*, le sens exact de cette dernière étant, du reste, douteux.

L'étrusque faisait aussi emploi de préfixes dont la valeur demeure encore inconnue [3].

Comme on le voit par cet exposé très succinct, nos connaissances grammaticales restent encore réduites. La lente exhumation et l'étude méthodique des langues d'Asie Mineure permettent, nonobstant, d'entrevoir une époque relativement prochaine où ce peuple cessera définitivement de mériter le qualificatif que lui

1. Ex. : *zilaχ* = fonction officielle mal définie > *zilaχce* = « il a exercé la charge zilakh » ; l'on trouve aussi le *-ce* précédé d'un infixe *-nu-* : *zil(a)χn(u)ce* (v. KÖRTE, dans PONTRANDOLFI, p. 150).

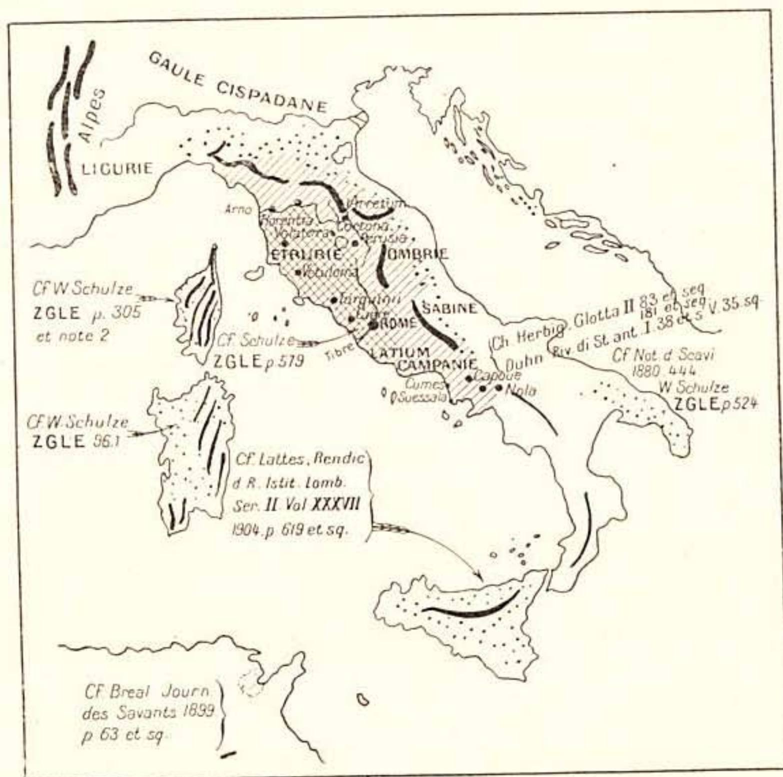
2. Cet enclitique se retrouve également en lydien ; v. LITTMANN, *Sardis*, index. Quant au *-m*, il paraît avoir sa contre-partie dans l'enclitique (?) hétéen *-ma* ; en tous cas, c'est une coïncidence à noter.

3. Il y a peu de bilingues étrusco-latins, et ils sont courts. Les principaux sont : CIE 3763 Pérouse = CIL XI, 1963 ; CIE 1671 Clusium = CIL 2660 ; CIE 378 Arrezzo = CIL 1855 ; CIL XI 6363 Pesaro, que voici : ..fatius L. F. Ste. haruspex fulguriator ; étrusque : *cafates lr. lr. netśvis trutnvt frontac* ; à traduire probablement : *L. Cafatius d'entrailles observateur* (?) *et de la foudre* (*fronta-c*) ; mais le mot à mot n'a pas encore été fourni.

décernait Denys d'Halicarnasse [1], de *peuple dont la langue ne lui est commune avec aucun autre.*

Si nous ne comprenons pas encore son idiome ; si les bandelettes de la momie d'Agram ne nous ont pas jusqu'ici livré leur

Croquis indiquant l'aire linguistique [2] de l'étrusque en Italie et la présence d'éléments étrusques dans le voisinage immédiat.



ZGLE = *Zur Geschichte Lateinischer Eigennamen*, Göttingue, 1904.

1. 28, I, 30.

2. L'aire linguistique propre est indiquée par des hachures croisées ; l'aire d'expansion approximative par des hachures simples ; la présence probable d'éléments diffus par un semis de points. La diffraction des éléments étrusques a été, en fait, beaucoup plus vaste, mais cette question dépasse notre cadre actuel.

secret ; si, des 8.000 inscriptions environ que comportera le *Corpus Inscr. Etrusc.*, nous ne pouvons saisir le sens que des plus courtes (qui nous valent quelques noms de parenté, tels que *clan* = fils, *sec* (ou *seχ*) [1] = fille, *puia* = épouse, *lautni* = affranchi [fém. *lautnita*], *avil* = aetas, *ril* = année, *tiv* = lune, *usil* = soleil, etc.), l'on peut du moins dire que les fouilles de l'Asie antérieure et des îles méditerranéennes orientales, en nous rendant des textes et des monuments ensevelis depuis bien des siècles, ont fait entrer la *glottologie* étrusque dans une phase nouvelle : celle de la *philologie comparée*. Mais l'on ne saurait trop insister sur ce fait que nous n'en sommes encore qu'aux tout premiers débuts [2].

Avec l'étrusque, la brève inscription de Lemnos [3] témoigne d'affinités assez remarquables, surtout en ce qui concerne, semble-t-il, les noms de nombre. Mais ce document est trop court pour prêter ici à des conclusions de large portée.

La LANGUE DES TABLETTES CRÉTOISES [4], écrites en caractères hiéroglyphiques plus ou moins cursifs, suivant les époques ou les scribes, n'a pas encore d'existence linguistique. Il faut attendre pour cela qu'on puisse la lire.

Quant aux idiomes encore inconnus retrouvés à CHYPRE [5] et

1. L'on trouve aussi *śec* et *śeχ*.

2. A l'étrusque, ou, du moins, au même groupe d'idiomes, il y a lieu de rattacher le rhétique, attesté seulement par des noms propres. On trouvera l'essentiel de la bibliographie du rhétique dans le *Dictionnaire des antiquités* de DAREMBERG et SAGLIO, sous *Étrusques*, p. 819, col. 1, notes ; dans le tome I des *Altitalische Forschungen* de PAULI et dans l'ouvrage de G. PONTRANDOLFI (voir *Bibliographie*, § ÉTRUSQUE, à la fin du présent chapitre).

3. L'essentiel de la bibliographie relative à l'inscription de Lemnos se trouve dans le livre précité de PONTRANDOLFI

4. Reproduites en partie par EVANS, dans ses *Scripta Minoa*, t. I. Londres. Les plus récentes études sont celles de J. SUNDWALL, dans les *Acta Academiae Aboensis Humaniora*, I, 2 et II, 3, 1919 et 1920.

5. Voir J. VENDRYES, *Inscriptions cypriotes en langue inconnue* (*MSL.*, t. XVIII).

à PRAISOS [1], ce sont des épaves insituables, pour l'instant, de la vaste et complexe civilisation méditerranéenne préhellénique. Le sens de ces inscriptions, courtes et mutilées, nous échappe et il suffit, sans doute, de les avoir mentionnées.

L'on en doit dire autant de la langue que recouvrent les hiéroglyphes du disque de Phaestos [2].

*
* *

Comme on le voit, l'histoire des langues examinées dans ce chapitre — du moins des plus importantes d'entre elles — n'est que l'un des aspects d'une étude singulièrement plus large et plus diverse : celle de l'expansion économique, militaire, politique, religieuse, puis du déclin des peuples qui les ont parlées.

A ce titre, elle se rattache directement à la connaissance d'époques lointaines, de régions imparfaitement explorées jusqu'ici, de groupes ethniques souvent assez mal définis, que leurs destinées ont, en outre, enchevêtrés au delà de toute expression [3]. Tout démontre, néanmoins, qu'ils ont dans leur ensemble exercé une action décisive sur l'élaboration des grandes civilisations classiques, orientales et méditerranéennes, et, par ce dernier intermédiaire, sur toute la culture ancienne de l'Occident.

Une grande religion, vaste théosophie de la nature, qui asso-

1. D. COMPARETTI, *Mus. Ital.*, II (1888), p. 673 et suiv. *Iscrizioni di varie Città Cretesi*; A. J. EVANS, J. H. S., vol. XIV, pp. 354 et suiv. *Primitive Pictographs and Scripts from Crete and the Peloponnese*; R. S. CONWAY, *B. S. Ath.* (1901-1902), pp. 125-6 ; et *ibid.*, 1903-4, pp. 115-7. L'on pourra se référer également à l'*Einleitung* de GERCKE et NORDEN [2], vol. I, p. 147, et à H. Th. BOSSERT, *Altkreta*, Berlin, 1921, p. 65-6, ainsi qu'aux *Scripta Minoa*, I, d'A. J. EVANS.

2. EVANS, *Scripta minoa*, p. 287. Le dernier article paru sur ce monument encore indéchiffré, est celui de A. ROWE, *The Phaestos disk : its Cypriote origin* (Transact. of the Roy. Soc. of South Austr., 1919 ; XLIII). Tant qu'elle n'aura point été lue, cette inscription n'offre point de prise aux linguistes.

3. C'est pourquoi le croquis donné ne porte pas d'indications de frontières linguistiques, l'onomastique démontrant des superpositions très complexes.

ciait les destinées de l'homme et de son groupe au renouveau annuel de la vitalité et de la fécondité terrestres ; une astrologie développée, une savante doctrine des présages (que les Étrusques importeront en Italie) avec les sciences d'application qui en procèdent : astronomie, droit, médecine, navigation au long cours; des écritures très usitées : la cunéiforme, la cappadocienne hiéroglyphique et hiératique (= syllabaire cypriote) ; une tendance très marquée vers l'alphabétisme, qui devait se traduire quelque jour par la constitution d'alphabets ; une curieuse activité législative et juridique ; le dressage du cheval, l'art de la fortification et des aménagements maritimes, la construction des navires, la métallurgie, l'agriculture (des céréales et des légumineuses, et surtout de la vigne), la fabrication et l'usage de l'huile et du vin ; telles sont les principales manifestations de l'activité propre de l'Asie Mineure.

Ces brèves considérations, d'ordre non exclusivement linguistique, il est vrai — mais la linguistique elle-même n'est que l'une des formes de l'histoire — aideront peut-être à comprendre l'extrême complexité des problèmes qui se posent ; la nature multiple des rapports ayant uni ces peuples aux habitants des contrées environnantes, où se parlaient de tout autres langues ; les liens ténus et subtils, réels cependant, qui les relient à des pays relativement proches, comme la Grèce, l'Égypte, ou déjà plus éloignés, comme l'Italie et les côtes méditerranéennes; la variété des influences exercées et subies.

Parmi ces actions et réactions mutuelles, il en est, cependant, de fondamentales que l'on ne saurait se dispenser de mentionner en terminant.

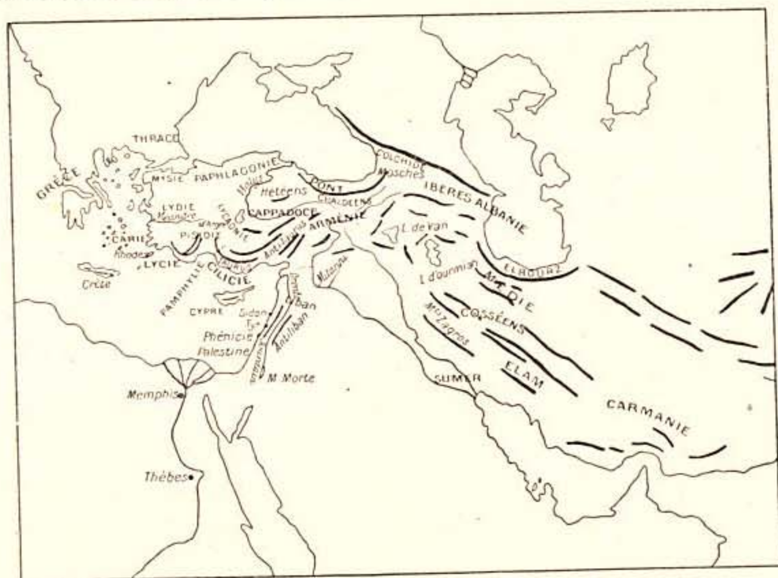
L'une est l'absorption permanente et, pour ainsi dire, la progressive *digestion* de l'asianisme par le sémitisme, dans toute la zone qui, *lato sensu*, s'étend du Sud du Taurus aux premières pentes de l'Irak.

Cette digestion, cela va sans dire, n'a pas eu lieu sans échanges réciproques nombreux.

L'on peut affirmer, néanmoins, que, sur cette aire géographique, tôt ou tard les idiomes de la famille sémitique finissent par prévaloir.

De leur lente revanche contre les éléments allogènes qui les ont passagèrement entamés, la sémitisation progressive du sumérien, les infiltrations sémitiques en élamite [1], le dialecte sémitique bâtard des tablettes cappadociennes, le complexe, encore mal

Croquis indiquant l'aire linguistique générale des idiomes asianiques et les zones d'expansion proches.



Les limites de chaque idiome n'ont, à dessein, pas été indiquées, mais seulement sa région, à cause des variations dans l'expansion de chacun. L'aire maxima a été signalée pour chaque langue au cours du paragraphe qui lui est consacré, lorsqu'il a été possible.

défini mais comportant en tous cas, une proportion *x* de sémitique, désigné jusqu'ici sous le nom d'AMORITE (OU AMORRHÉEN) [2] constituent autant de témoignages irrécusables.

1. Voir G. HÜSING, *Semitische Lehnwörter im Elamischen* (Beitr. z. Assyr. 5, p. 405-412).

2. Au sujet des Amorites, le dernier ouvrage d'ensemble est celui de A. T. CLAY, *The Empire of the Amorites*, New Haven-Londres, 1919 (vol. VI des Yale Oriental series : researches) ; à n'utiliser qu'avec précaution. Les Amorites n'ont, jusqu'ici, survécu que par des noms propres qui ne sont, d'ailleurs, pas tous sémitiques.

Une autre action non moins incontestable, plus délicate cependant à déceler, est celle exercée par les parlers asianiques sur les vocabulaires de l'Égypte, de la Grèce, de l'Italie à la faveur du négoce et des entreprises coloniales ou militaires des clans anatoliens et méditerranéens [1].

Il ne saurait entrer, dans le cadre de ce bref chapitre, d'examiner, fût-ce sommairement, cet aspect de la question, mais il était nécessaire d'y faire, à tout le moins, une allusion.

*
* *

Après une période d'expansion que l'on peut situer antérieurement au Ve millénaire pour le sumérien et, pour les autres idiomes, entre 4.000 et 1.500 av. J.-C., ces parlers ont donc peu à peu reculé devant le sémitique. Puis, vers la moitié du IIe millénaire, l'influence indo-iranienne se manifeste clairement par des emprunts théologiques, par des noms de princes, des contaminations lexicologiques. L'influence indo-européenne s'accentuera lors de la poussée arméno-phrygienne vers l'Est et vers le Sud.

La « guerre de Troie », la fondation de l'Empire perse, l'expansion du monde hellénique porteront, à leur tour, de nouveaux coups à ce qui reste encore des idiomes asianiques.

Ces derniers survivront, cependant, par leurs tendances phonétiques, par d'innombrables vestiges de vocabulaires subsistant en sémitique, en indo-iranien, en arménien [2], en grec, etc.

1. Voir A. MEILLET, *M. S. L.*, t. XV, pp. 161-164 ; A. CUNY, *Rev. des Et. anc.*, 1910, pp. 154-164 ; C. AUTRAN, *Phéniciens*, pp. 5-6 et 45-47, A. MEILLET, *Aperçu d'une histoire de la langue grecque*, 2e édit. Paris, 1920, chap. III et la Bibliographie, en tête du volume, ainsi que R. DUSSAUD, *Les Civilisations préhelléniques*, etc., 2e édit. Paris, 1914, pp. 437-442. Voir aussi les dictionnaires étymologiques grecs de BOISACQ et de F. MÜLLER (ce dernier, dictionnaire général, mais comportant une partie étymologique excellente) et latin de A. WALDE, 2e édit. Heidelberg, 1910. Pour les mots ÉGÉENS en égyptien, voir AUTRAN, *Tarkondemos*, ch. III, § 1.

2. A. MEILLET, dans son *Esquisse d'une grammaire comparée de l'arménien classique*, Vienne, 1903 (début), a montré comment, par une influence analogue, l'action des langues caucasiennes s'est manifestée dans la formation de la grammaire arménienne.

Vestiges que notre ignorance nous empêche, malheureusement, de reconnaître dans la plupart des cas, mais qui, sans aucun doute, suffisent à rendre en partie raison de la forte proportion de termes d'origine inconnue que l'on relève dans ces langues.

Avec la conquête d'Alexandre, c'en est fait de tout ce vieux monde.

L'hellénisme vainqueur étend désormais son voile uniforme sur toute l'Asie Occidentale, du Pont à l'Indus. Arts, costumes, science, langue officielle, tout devient grec — en apparence, du moins

Chose infiniment plus grave : ce qui n'est pas hellénique passe pour barbarie, est négligé, méprisé, détruit par cette *mode* mille fois plus intolérante que ne l'avait jamais été la monarchie achéménide, si paternelle, au fond.

A cet effacement quasi total des anciennes civilisations indigènes au profit d'une culture envahissante, la linguistique, l'histoire politique, celles de l'art, des religions, de l'esprit humain doivent d'irréparables pertes que les fouilles entreprises en Anatolie, en Mésopotamie, en Susiane, en Lydie, en Palestine permettront, il faut l'espérer, de réparer en partie.

C. Autran.

BIBLIOGRAPHIE

Périodiques renfermant l'essentiel des travaux, qui sont très dispersés.

Allemagne et Autriche. — Zeitschrift der Deutschen Morgenländischen Gesellschaft, Orientalistische Literaturzeitung, Wiener Zeitschrift für die Kunde des Morgenlandes, Zeitschrift für Assyriologie, Bezzenbergers Beiträge, Kuhns Zeitschrift, Indo-germanische Forschungen, Abhandlungen et Berichte des Académies de Berlin, Vienne, Münich, de l'Académie de Saxe ; les revues Glotta et Memnon, Athenische Mitteilungen, Jahrbuch des Archäologischen Instituts in Athen, Mitteilungen der Deutschen Orient-Gesellschaft, der Vorderasiatischen Gesellschaft, Zeitschrift für Ægyptische Sprache und Altertumskunde, Anthropos, Zeitschrift für Indologie und Iranistik, Orientalische Bibliographie, Archiv für Religionswissenschaft, Zeitschrift der Alttestamentlichen Wissenschaft, Abhandlungen der K. Gesellschaft der Wissenschaften

(Göttingen), Mitteilungen des Deutschen Pälestia-Vereins, Wörter und Sachen, Zeitschrift für Semitistik und verwandte Gebiete, Zeitschrift für Numismatik, Anatole, Klio, etc.

Finlande. — Finnisch-ugrische Forschungen — Journal de la Société finno-ougrienne.

Amérique et Angleterre. — American Journal of Archaelogy, American Journal of semitic languages and literatures, Annual of the British School of Athens, Publications de l'Egypt exploration fund et du Palestine exploration fund, Journal of Hellenic Studies, Journal of the royal Anthropological Institute (Liverpool), University annals of Art and Archaeology, Proceedings and transactions of the Society of biblical archaeology, Proceedings of the British Academy of the Royal Geographical Society, of the Society of Antiquaries, Classical review, American journal of philology, Memoirs of the American academy of arts and sciences (aussi Proceedings), Philologica, Bulletine of the School of Oriental Studies, Journal of the American Oriental society, Journal of the royal Asiatic Society, etc.

France et Belgique. — Annales du service des Antiquités d'Égypte, Mémoires de l'Institut français au Caire, Mémoire de l'Institut Égyptien, Recueil de travaux relatifs à l'archéologie et à la philologie égyptiennes et assyriennes, Mémoires de la société des Antiquaires français, Revue Égyptologique, Revue d'Assyriologie, Revue biblique, Revue Syria, Revue internationale des études basques, Mémoires et Bulletin de la Société de linguistique de Paris, Journal asiatique, Revue des études grecques, Bulletin de correspondance hellénique, Revue des études juives, Revue archéologique, Revue de numismatique, Muséon, Publications de l'Académie royale de Belgique, Revue de Philologie, Revue des Études anciennes, Hespéris, etc.

Italie. — Notizie degli scavi..., Ausonia, Bolletino di Paletnologia italiana, Ægyptus, Archivio storico italiano, Atene e Roma, Studi italiani di filologia classica, Archivio glottologico italiano, Rivista indo-greco-italica, Nuova antologia, Rendiconti del reale Istituto Lombardo, Memorie della reale Accademia di Napoli, Publications de l'Accademia dei Lincei, etc.

Grèce. — Ephemeris Archaiologiké.

Sumérien. — C. Fossey, *Manuel d'Assyriologie*, t. I (seul paru), Paris, 1904, utile pour l'historique de la « question sumérienne ».

St. Langdon, *A Sumerian grammar and Chrestomathy, with a vocabulary of the principal roots in Sumerian and a list of the most important syllabic and vowel transcriptions*, Paris, 1911. Ouvrage sérieux et bien construit, mais qu'il est nécessaire de compléter par :

Frdr. Delitzsch, *Grundzüge der sumerischen Grammatik*, dans la collection des Hilfsbücher zur Kunde des Alten Orients (Band 5), Leipzig, 1914.

Les personnes non au courant de la philologie assyrienne pourront avoir recours à la : *Kleine sumerische Sprachlehre für Nichtassyriologen* (sommaire

grammatical, vocabulaire, spécimens de textes), du même, Leipzig, 1914, et à B. MEISSNER, *Die Keilschrift*, 2e édit., 1922 (collection Göschen, n. 708).

Ces deux ouvrages constituent les études grammaticales d'ensemble les plus récentes. L'on y trouvera tous les éléments d'une bibliographie plus étendue avec l'indication des collections de textes.

Comme lexiques, les plus récents et les meilleurs sont ceux de :

A. DEIMEL, *Vocabularium sumericum ad textus archaicos*, Rome, Instit. bibl. pontif., 1910 ;

FRDR. DELITZSCH, *Sumerisches Glossar*, Leipzig, 1914 ; ce qu'il y a de plus complet, pour le moment, comme manuel lexicographique en caractères latins.

Les travaux nombreux de F. THUREAU-DANGIN sont fondamentaux.

L'on trouvera dans le Manuel bibliographique intitulé *Die Assyriologie 1914-1922*, publié par E. F. WEIDNER, Leipzig, Hinrichs, 1922, l'indication des derniers travaux parus jusqu'au 31 juillet 1922 touchant le sumérien.

Mitannien. — Texte : publié d'abord par :

WINCKLER et ABEL, dans les *Mittheilungen aus den oriental. Sammlungen der Kgl. Museen in Berlin*, Heft I (no 27) ; ensuite par :

J. A. KNUDZTON, dans les *Beiträge zur Assyriologie*, etc., IV, pp. 134-153 (en transcription) ;

La première discussion linguistique d'ensemble, avec essai de traduction et lexique a été donnée par :

L. MESSERSCHMIDT, *Mitanni-Studien*, dans les *Mitteil. der Vorderas. Gesellschaft*, 1899, Heft 4 ; 134 pp. 8o ;

Voir également : F. BORK, *Die Mitannisprache*, M.V.A.G., 1909, 1-2, qui paraît avoir mieux pénétré que Messerschmidt dans le détail du fonctionnement grammatical.

A. GUSTAV, *Verbindungslinien zwischen dem Mitannischen, dem Elamischen und dem Lykischen* (Memnon., VII, 4, 1915, pp. 228-232).

V. SCHEIL, *Tablettes de Kerkouk*, *Revue d'Assyriologie*, XV, 2, 1918, pp. 65-73 ; noms propres mitanniens.

Aussi E. F. WEIDNER, *Assyriologie*, p. 188.

J. A. KNUDZTON, O. WEBER, E. EBELING, dans leurs EL-AMARNA-TAFELN, Leipzig, 1915 (2 vol.), ont fait ou suggéré un certain nombre de corrections (voir t. I, p. 180 et suiv. ; lettre 24).

Pour la littérature jusqu'en 1913, voir la *Geschichte des Altertums* d'EDOUARD MEYER, 3e édit. Stuttgart et Berlin, t. I, 2e moitié. L'Orientalistische Literaturzeitung et Memnon renferment, pour les années subséquentes, l'essentiel des travaux parus.

V. aussi E. F. WEIDNER, *Assyriologie*, pp. 122-125.

Cosséen. — FRDR. DELITZSCH, *Die Sprache der Kossäer*, Leipzig, 1884, demeure encore le travail fondamental.

J. SCHEFTELLOWITZ, *die Sprache der Kossäer* (*Kuhn's Zeitschrift*, 1905, pp. 260-277), témoigne d'un faux point de vue linguistique (systématiquement indo-européen).

L'on pourra également consulter des études de F. BORK, dans *Memnon*, V, pp. 45 et suiv.; *Orient. Literat. Ztg.*, IX, 1906, pp. 588 et suiv.; *Beiträge zur Sprachwiss.*, III, 13 et suiv.;

GUSTAV, O.L.Z., 1912, p. 214 et suiv.; 300 et suiv.; 350 et suiv., etc., ainsi que les périodiques usuels relatifs à l'ancien Orient.

A. T. CLAY, *Personal names from cuneiform inscriptions of the Cassite period* (Yale Oriental series, vol. I), New Haven, Londres, Oxford, 1912, notamment pp. 36-45, renferme l'essentiel au point de vue de l'onomastique.

G. HÜSING, dans *Memnon*, IV, 1910, pp. 24 et suiv., a exprimé à l'égard du court lexique cosséen-sémitique reproduit par FRDR. DELITZSCH des doutes qui semblent excessifs.

T. G. PINCHES, *Journ. of the Royal Asiatic Society*, 1917, p. 101 et suiv., complète DELITZSCH sur certains points.

Pour les ressemblances entre le peu de cosséen connu et les langues caucasiennes, voir F. BORK, O.L.Z., 1911, pp. 472 et suiv.

Aussi E. F. WEIDNER, *Assyriologie*, p. 188.

Vannique. — Les inscriptions de ce langage ont été réunies par : Jos. SANDALGIAN, *Les inscriptions cunéiformes urartiques, etc. Venise. Mékhitaristes, 1900.* Publication faite sans critique et sous l'influence d'idées préconçues (l'auteur est arménien). Elle ne dispense donc pas de recourir aux travaux antérieurs parmi lesquels les principaux sont :

A. H. SAYCE, *The cuneiform inscriptions of Van*, etc., dans le *Journ. of the Roy. asiat. Society*, XIV (1882), pp. 377 et suiv., et 732 (résultats tout provisoires); BELCK, *Die Kelischinstele*, dans *Anatole*, I (1904) ; KLUGE, dans les *Mitteil. der Vorderas. Ges.*, 1907, fasc. 5 ; WEIDNER, *Assyriologie*, p. 178, et la nouvelle édition des textes : MARR et ORBELI, Arkheologičeskaja Ekspedicja, 1916, etc., Pétrograd, 1922, VIII-68 pp., 7 fac-similés, 20 pl.

Élamite. — Les textes anciens ont été publiés par le P. V. SCHEIL dans les *Mémoires de la Délégation française* en Perse publiés sous la direction de J. DE MORGAN à partir de 1900 : tomes II, IV, VI, X. Textes élamites-sémitiques; tomes III, V, IX, XI. Textes élamites-anzanites.

F. H. WEISSBACH, *Anzanische Inschriften* et *Neue Beiträge*, etc., parus dans les *Abhandl. der sächs. Gesellsch. der Wiss.*, XII (1891), pp. 115-150, et t. XIV (1894), pp. 729-777, et 5 Pl.; voir aussi FOY, *Zeitschr. d. d. Morg. Ges.*, 52 et 54.

Les textes en néo-susien se trouvent dans :

F. H. WEISSBACH, *Die Achämenideninschriften zweiter Art* (Assyriol. Bibliothek, vol. IX). Leipzig, 1890 (texte, transcription, syllabaire, rudiment de grammaire et lexique, avec introduction et bibliographie) ;

Du même : *Die Keilinschriften der Achämeniden* (Vorderas. Biblioth.),

Leipzig, 1911 (transcription et traduction des inscriptions en v. p., él. et babyl. des rois achéménides), où l'on trouvera un supplément de bibliographie).

Consulter en outre :

FRDR. DELITZSCH, *Die Sprache der Kossäer* ;

G. HÜSING, *Elamische Studien* (M.V.A.G., 3, 7), 1898, nº I; seul paru (monographie de 42 pp. et 1 Pl.).

Du même : *Die einheimischen Quellen zur Geschichte Elams* (Assyriol. Bibl. 24, 1), Leipzig, Ire Partie, 1916 ; très utiles l'un et l'autre tant au point de vue philologique que bibliographique.

F. BORK, *Zur Erklärung der Elamischen Briefe* (Beiträge zur Assyr., 5, 3, pp. 401-404) ; voir aussi sa recension du livre précédent de Hüsing, dans l'*Orient. Liter. Ztg.*, XX (1917), pp. 79-83.

Le plan de travail de la *Vorderasiatische Bibliothek* (Leipzig) comprend un volume de *Textes élamites* (vol. 14 projeté), dont la date de publication ainsi que l'auteur ne peuvent encore être indiqués. — L'on fera bien également de se référer aux *Iranische Eigennamen in den Achämeniden Inschriften* de G. HÜSING, Dissert. 1907 ; *du même*, v. les articles portant le titre commun de *Kaspisches*, OLZ, XX, pp. 106-109 (sur la déesse *Šumalija*) ; OLZ, XX, pp. 178-181, sur *DURU-Katli-me* (étude d'onomast. géogr.) ; pp. 205-209 sur le verbe élamite.

Aussi WEIDNER, *Assyriologie*, pp. 178-180, et CRUVEILHIER, Principaux résultats des nouvelles fouilles de Suse, Paris, 1921.

Hétéen. — A. *Monuments hiéroglyphiques.* Rien ne prouve jusqu'ici que cette écriture recouvre spécialement du hétéen. L'on persiste, néanmoins, à la désigner par le terme d'hiéroglyphes hétéens; pure convention.

Les tentatives de déchiffrement les plus récentes sont celles de :

R. C. THOMPSON, *A new decipherment of the Hittite hieroglyphics*, Oxford, 1913. Pas très convaincant.

A. E. COWLEY, *The Hittites*, Schweich Lectures, 1918. Londres, 1920 ; mérite beaucoup de réserves, malgré la science et l'ingéniosité de l'auteur ;

C. AUTRAN, *Tarkondemos, réflexions sur quelques éléments graphiques figurant sur le monument appelé sceau de Tarkondemos*, Paris, 1921-1922, est un essai de discussion critique sur la valeur de deux signes et sur les conséquences historiques et linguistiques en résultant : nullement une tentative de lecture.

CARL FRANK, *Die sogenannten hettitischen Hieroglypheninschriften*, ein neuer Beitrag zu ihrer Entzifferung, *Abhandlungen für die Kunde des Morgenlandes*, XVI, 3 (1923).

Les textes ont été réunis par :

L. MESSERSCHMIDT, *Corpus inscr. hettiticarum*, 4 fasc. 1900-1906, Leipzig.

Le BRITISH MUSEUM a commencé à publier en 1914 un certain nombre de photographies d'inscriptions dans *Carchemish*, Part I (D. G. HOGARTH).

B. *Monuments en écriture cunéiforme.*

FRDR. DELITZSCH, *Sumerisch-Akkadisch-Hettitische Vokabularfragmente* dans les *Abh. der Kgl. preuss. Ak. d. Wiss. (Phil. Hist. Classe)*, 1914, 3 ;

DEUTSCHE ORIENT-GESELLSCHAFT : *Keilschrifttexte aus Boghazkoi* ;

Fasc. I, par H. H. FIGULLA et E. F. WEIDNER, Leipzig, 1916;

Fasc. II, par H. H. FIGULLA, 1916 ;

Fasc. III, par H.H. FIGULLA et O. WEBER, 1919 (1re moitié) ;

Fasc. IV, par E. FORRER, 1920;

Fasc. V et VI, par Fr. HROZNÝ, 1921.

Die Boghazkõi-Texte in Umschrift.

I. Band. *Einleitung, Die Keilschrift von Boghazkõi*, par E. FORRER, Leipzig, 1922.

II. Band, Fasc. 1. *Geschichtliche Texte aus dem alten Chatti-Reich*, par E. FORRER, 1922.

Frédéric HROZNÝ, *Code hittite, 1re partie*, transcription, traduction française. Paris, 1922.

Le même auteur prépare la publication d'un traité de même origine relatif au dressage du cheval.

Pour l'essentiel de la bibliographie, voir :

E. F. WEIDNER, *Assyriologie*, pp. 180-187.

G. CONTENAU, *Éléments de bibliographie hittite*, Paris, 1922.

Voir en outre : E. FORRER, *Die Inschriften und Sprachen des Hattischen Reiches*, Ztsch. d. d. Morgenl. Gesellsch. Band I, Neue Folge, 1922, et, sur la langue en particulier, FRIEDRICH, *Die hethitische Sprache* à la suite, dans le même recueil.

HERBIG, dans *Indogermanisches Jahrbuch*, VIII, p. 1-20.

C. *Arzawa. Yuzgat. Eurdek burnu.*

A la série des monuments rédigés en « hétéen » cunéiforme, il y a lieu de joindre :

1° Un certain nombre de tablettes ou fragments de même type linguistique récoltés et publiés par E. CHANTRE (*Mission en Cappadoce*, Paris, 1898, p. 40 et suiv.) ; omises par Delitzsch dans sa bibliographie ;

2° Les tablettes d'Arzawa (Collection de Tell el-Amarna) ; = lettres 31 et 32 de la publication de KNUDZTON, O. WEBER et E. EBELING, Leipzig, 1915 [texte seul en transcription] ; édition du texte avec essai de traduction par :

J. A. KNUDZTON, *Die Zwei Arzawa-Briefe*, Leipzig, 1902, avec remarques de S. BUGGE et d'A. TORP. — L'ouvrage de Knudzton, Weber et Ebeling renferme de nombreuses rectifications à ce travail ; voir également O. SCHRÖDER, *Orient. Liter. Ztg.*, 1915, p. 231 et sq. ; KNUDZTON, *ibid.*, XIX, pp. 135-137 et SCHRÖDER, *ibid.*, p. 138.

3° Les tablettes de Yuzgat : A. H. SAYCE et T. G. PINCHES, Roy. As. Soc. Mon., vol. XI, Londres, 1907 ; A. H. SAYCE, *Journ. of the roy. Asiat. Soc.*, 1909, pp. 963-980, s'y réfère aussi. *Liverpool Annals*, III (1910), pp. 99-106 ;

Enfin A. H. SAYCE, J. R.A.S., 1920, 1, pp. 70-83, dernier essai de traduction de la plus importante de ces tablettes ;

4° Inscription — en caractères araméens — d'Eurdek burnu (près Zendschirli). Voir LIDZBARSKI, *Ephemeris*, III, p. 192.

Voir aussi E. F. Weidner, *Assyriologie*, pp. 119-125.
Pour l'Amorite, *ibid.*, pp. 176-7.

Carien. — La plupart des inscriptions ont été publiées par Sayce dans les *Transact. Soc. Bibl. Arch.* (1893), pp. 112-154, 4 Pl.; *Proc. Soc. Bibl. Arch.*, XVII (1895), pp. 39-43 ; XXVII (1905), pp. 123-128, 2 Pl. ; XXX (1908), p. 28, 1 Pl. ; XXXII (1910), pp. 261-2, 1 Pl. Elles ont été réunies et republiées avec quelques autres par Sundwall, dans *Klio*, 1911 (pp. 26 et suiv., 464 et suiv.). Le reste du matériel carien (gloses et noms propres) a été rassemblé, sans grande critique, par G. Meyer, dans les *Bezzenberger Beiträge*, t. X (1886), pp. 147-202. L'on trouvera quelques indications utiles dans P. Kretschmer, *Einleitung in die Geschichte der Griechischen Sprache*, Göttingen, 1896, pp. 376-394, ainsi que dans les premiers volumes de l'*Einleitung in die Altertumswissenschaft* de Gercke et Norden. Voir également G. Dottin, *Anciens peuples de l'Europe*, Paris, 1916, pp. 113-116 ; C. Autran, *Phéniciens*, pp. 22 et ss. L'auteur a réuni les matériaux nécessaires à un travail d'ensemble sur la question, à paraître.

Lydien. — On trouvera dans E. Littmann, *Sardis*, Leyde, 1916, des inscriptions lydiennes découvertes par l'expédition américaine ainsi que l'essentiel de la bibliographie. Consulter en outre les ouvrages de Kretschmer et de Dottin cité pour le carien, ainsi que les périodiques énumérés en tête de la présente bibliographie. Parmi les derniers articles utiles, signalons celui de R. Thurneysen, *Zum Lydischen*, Kuhns Ztschr. Bd. 50, Janv. 1922.

Lycien. — Inscriptions et bibliographie dans E. Kalinka, *Tituli Asiae Minoris*, I, Vienne, 1901. Voir également les ouvrages de Kretschmer et Dottin, pour les généralités. Les noms propres ont été recueillis par Sundwall dans *Klio*, 1913 (XI Beiheft). L'on pourra aussi consulter Th. Kluge, *die Lykier* (der Alte Orient 1910, 11, 2), brochure de vulgarisation de 32 pp., C. Autran, *Tarkondemos*, chap. III, donne quelques détails d'interprétation complémentaires. Voir aussi l'article de H. Pedersen sur les inscriptions lépontiennes dans la revue *Philologica*, 1921.

Pour toutes ces langues l'on glanera çà et là dans E. F. Weidner, *Assyriologie* des informations utiles.

Étrusque. — Tout l'essentiel, comme information et bibliographie se trouve dans G. Pontrandolfi, *Gli Etruschi e la loro lingua*, Florence, 1909 (traduction et remise au courant, avec concours des auteurs, des articles de Körte : *Etrusker* ; de Thulin : *Etrusca disciplina*, de Skutsch, *Etruskische Sprache*, parus dans l'Encycl. de Pauly-Wissowa. La section consacrée à la langue occupe les pp. 95-164. Le *Katalog der Bibliothek des Kais.-Deutsch. Archäol. Instit. in Rom*, par A. Mau, refait en 1913-1914 (Rome), par E. von Mercklein, ren-

ferme un supplément bibliographique utile : de même la *Topografia storica dell Etruria* d'ART-SOLARI (vol. I, Pise, 1918 ; II, 1914 ; III, 1915 ; IV, 1915).

MARTELLO GINO LUIGI a publié à Pérouse, 1920, sous le titre : *Lingua etrusca : grammatica, testi con traduzione a fronte, glossario*, un essai grammatical d'une centaine de pages dont une partie n'apporte que ce que l'on savait déjà et l'autre demeure très incertaine (les interprétations personnelles de l'auteur). Les tentatives d'explication de J. MARTHA, *La langue étrusque*, Paris, 1913 ; du P. HILAIRE DE BARENTON, *La langue étrusque, dialecte de l'Ancien égyptien*, Paris, 1920, ne méritent pas d'être retenues. Pas davantage les *Tusca* d'A. GRÜNWEDEL, pure fantaisie.

L'ouvrage fondamental reste, MÜLLER-DEECKE, *Die Etrusker*, Stuttgart, 1877.

G. HERBIG, qui prépare avec le gracieux concours de B. NOGARA, une « synthèse étrusque », à paraître, donne en tête de son article *Etruskisches Latein* dans les *Indo-germaniche Forsch.*, 1916-7 (t. XXXVII), une bibliographie de sources modernes utiles. Voir aussi G. PINZA, *Storia della civiltà latina*, etc., en voie de publication très prochaine. L'on trouvera quelques indications sur des points de phonétique étrusque ainsi que sur quelques-uns des éléments constitutifs de ce groupe ethnique dans C. AUTRAN, *Tarkondemos*, chap. III.

Les gloses anciennes ont été colligées dans l'ouvrage de PONTRANDOLFI ; les inscriptions sont publiées : n[os] 1-4917 dans le t. I. du *Corp. inscr. etrusc.*, Leipzig, 1893-1902, par O. A. DANIELSSON et C. PAULI ; la publication du t. II (Sect. I, fasc. 1 : 4918-5210 ; sect. II, fasc. 1 : 8001-8600 parues), assurée par les soins de O. A. DANIELSSON, G. HERBIG et B. NOGARA (1907 et 1912), est en cours. Les textes de la momie d'Agram ont été publiés pour la première fois in-extenso par J. KRALL, dans les *Mém. de l'Acad. de Vienne*, Phil.-Hist. Klasse, t. XLI, 3, 1892 (70 pp. et X Pl.). L'on en trouvera de larges extraits dans J. MARTHA, *Langue étrusque*. Voir également G. HERBIG, *Die etruskische Leinwandrolle des Agramer National-Museums*, Münich, die Akademie, 1911, in-4°, 45 pp. (Abh. der Kgl. bayer. Ak. der Wiss. Philos. phil. u. hist. Klasse, XXV, 4).

Le *Nuovo saggio sulla lingua etrusca*, etc., du prof. G. BUONAMICI, Faenza, 1910-1911, ne comporte encore que la description des caractères externes de la langue et l'étude de l'alphabet et de la phonétique (I, 1 et 2).

Sur l'ensemble des langues asianiques et sur leurs rapports possibles soit entre elles, soit avec les idiomes préclassiques de la Méditerranée, un essai compact et pas toujours sûr, mais utile, a été fourni par K. OŠTIR, *Beiträge zur alarodischen Sprachwissenschaft*, vol. I, Vienne et Leipzig, 1921.

L'on trouvera quelques indications éparses dans les périodiques signalés ci-dessus (articles relatifs soit à l'Asie Mineure asianique, soit aux langues modernes du Caucase), ainsi que dans E. F. WEIDNER, *Assyriologie*.

Domaine égéen. — G. GLOTZ, *La civilisation égéenne*, Paris, 1923 (pp. 439-443).

C. AUTRAN, *Introduction à l'étude critique du nom propre grec*, Paris, 1923.

LA LANGUE BASQUE

Le basque, appelé *euskera*, *euskara*, *eskuara*, etc. par ceux qui en font usage, est parlé, pour les quatre cinquièmes environ en Espagne, et pour un cinquième en France, au Nord et à l'Ouest des Pyrénées occidentales, c'est-à-dire dans les provinces de Biscaye, Alava, Guipúzcoa et Navarre d'une part, dans les arrondissements de Bayonne et de Mauléon de l'autre. Il n'occupe point la totalité des divisions administratives qui viennent d'être énumérées : en effet, la partie occidentale de la Biscaye, constituant à peu près le quart du territoire de la province, n'est pas basque quant à la langue, de même que les neuf dixièmes de l'Alava, et plus de la moitié de la Navarre. En France, le domaine de l'*euskera* correspond à peu de chose près aux limites des anciennes provinces de Labourd, Basse-Navarre et Soule qui, avec le Béarn, constituèrent le département des Basses-Pyrénées. En outre, le basque est encore parlé assez fréquemment en Amérique par quelques milliers d'« Euskariens » qui, ayant conservé le culte de leur langue maternelle, ont fondé au delà des mers des associations, des journaux et des revues basques.

A combien peut-on évaluer le nombre des gens qui se servent de cette langue ? Il y a des bilingues, en proportions variées, dans tous les villages basques : et, dans certains d'entre eux, le total des bascophones est très inférieur à celui des personnes qui ne se servent que d'un idiome roman. Il est donc très difficile de répondre à la question posée. Le prince Louis-Lucien Bonaparte, l'auteur qui a le plus minutieusement étudié le problème de la répartition de l'*euskera*, évaluait en 1873 à 660.000 le nombre

des Basques-espagnols n'ayant pas abandonné la langue de leurs pères, et à 140.000 celui des Basques-français se trouvant dans le même cas. Jusqu'à plus ample informé, il convient de s'en tenir à ces chiffres, sans oublier toutefois que dans la Navarre transpyrénéenne notamment, l'*euskera* a sensiblement reculé depuis 1873.

Ce recul n'a cessé de se produire d'une façon plus ou moins intense, à toutes les frontières du pays basque, depuis une époque très ancienne, qu'il est impossible de déterminer. Divers témoignages, en sus des données de la toponymie, nous permettent d'affirmer que le basque s'étendait autrefois bien au delà de ses limites actuelles, surtout en Espagne.

*
* *

Une chose frappe avant tout celui qui veut observer la langue basque, qu'il la prenne dans sa réalité vivante ou dans les livres, c'est son extrême diversité dialectale.

Plusieurs auteurs, à diverses époques, se sont occupés de classer ces dialectes. Nous dirons simplement un mot de la classification que le prince L.-L. Bonaparte proposa il y a une cinquantaine d'années. Il divisait le basque en trois groupes dialectaux, cinquante dialectes (plus quatre dialectes littéraires), vingt-cinq « sous-dialectes » se fragmentant en cinquante « variétés » donnant elles-mêmes lieu à une dizaine de « sous-variétés ». C'est surtout par l'étude du verbe que le Prince parvint à cette consciencieuse classification. Il suffirait peut-être de distinguer deux grands groupes dialectaux : le biscayen (que l'on pourrait appeler aussi basque occidental) d'un côté, et de l'autre côté tous les autres dialectes (guipuzcoan, dialectes de la Haute et de la Basse-Navarre, labourdin, souletin) (voir planche 7). On pourrait appeler ce groupe, par opposition au premier, centro-oriental. Nous justifierions cette classification par la considération suivante : on passe par gradations insensibles d'un parler à l'autre parmi ceux qui constituent ce groupe, tandis que le saut est brusque lorsqu'on passe du guipuzcoan au biscayen. Ce dernier dialecte se distingue

en effet, dans toute une partie de son verbe, par l'emploi d'auxiliaires qui lui sont propres, il offre des particularités qui n'appartiennent qu'à lui dans maints détails de la grammaire en plus grand nombre que les autres dialectes, et enfin son vocabulaire tranche aussi assez souvent sur celui de ses congénères.

Les divisions dialectales ne correspondent pas toujours aux divisions géographiques; par exemple, le biscayen n'est pas seulement parlé en Biscaye, mais dans les rares villages de l'Alava où le basque est encore conservé et enfin dans une partie notable du Guipuzcoa. Cette dernière province n'est pas exclusivement guipuzcoane de langue : le haut-navarrais septentrional l'envahit en effet par le Nord, mais en revanche, le guipuzcoan se parle dans une partie de la Navarre. De même pour le labourdin, dans le domaine géographique duquel le bas-navarrais occidental et même oriental se parle, est en revanche en usage, à peine différent, dans presque toute la vallée transpyrénéenne de Baztan, etc.

Il va sans dire d'ailleurs que toutes ces classifications n'ont qu'une valeur relative, car dans les parlers basques il existe une énorme quantité de lignes isoglosses indépendantes les unes des autres.

*
* *

Les noms de lieu sont les plus anciens documents basques incontestables : on en cite qui remontent au VIIIe siècle. Les phrases ne commencent à paraître que beaucoup plus tard, et nous ne trouvons de textes de quelque étendue qu'à partir du XVIe siècle. Le premier livre est un recueil de poésies *Lingvæ Vasconvm primitiæ*, œuvre d'un prêtre, Dechepare : il fut édité à Bordeaux en 1545. Depuis cette époque, la littérature basque se compose principalement de traductions d'ouvrages religieux. Cependant on peut y joindre, en très petit nombre, quelques ouvrages originaux. Mais ce n'est qu'à partir de 1880 environ que ces derniers devinrent plus nombreux que les traductions. Vers cette époque, une véritable renaissance basque, plus ou moins liée à des théories politiques, s'est produite, et aujourd'hui il paraît

assez régulièrement des romans, des poèmes, des articles de journaux variés. A cela, il faut joindre la littérature orale, contes et chansons en grand nombre que l'on recueille depuis quelques années, et les « pastorales » souletines, pièces de théâtre dont les sujets, la facture et la technique sont empruntés, dans ce qu'ils ont d'essentiel, aux mystères français du moyen âge. Signalons enfin le don, qui se rencontre ailleurs, mais que les Basques possèdent particulièrement, des improvisations poétiques chantées.

*
* *

Les origines de la langue basque ont préoccupé beaucoup d'érudits et on n'en est plus à compter les hypothèses souvent aventureuses auxquelles elles ont donné lieu. La plus sérieuse de ces hypothèses paraît être celle, déjà ancienne, d'après laquelle le basque serait le dernier vestige de la langue des Ibères, ou d'une des langues parlées par les Ibères.

L'état linguistique de l'Ibérie nous est fort mal connu. En dehors de quelques mots et noms de lieu que nous ont transmis les écrivains grecs et latins, nous avons des médailles et des inscriptions quelques-unes celtibériennes (de temps en temps on en découvre de nouvelles). Ces inscriptions et les légendes des médailles sont écrites dans un alphabet qui ressemble à l'alphabet phénicien. On les déchiffre de diverses façons : c'est ainsi que la fameuse lame de plomb de Castellon nous donne un texte qui a été l'objet de cinq ou six interprétations n'ayant absolument rien de commun entre elles. Néanmoins, les quelque deux cents noms de personnes et de divinités qui nous restent de l'Aquitain ont une physionomie nettement basque (voir Luchaire, *Origines linguistiques de l'Aquitaine*, Paris, 1879) et il est probable que l'*euskera* actuel descend en ligne directe de ce dialecte ibérien.

Parmi les autres rapprochements qui ont été faits entre le basque et divers idiomes, ceux qui signalent des concordances avec des langues chamito-sémitiques méritent sans doute une certaine considération. D'autre part plusieurs linguistes supposent que le basque appartiendrait à une grande famille méditerranéenne, dont feraient partie les langues caucasiques, l'étrusque, etc.

*
* *

Ce qui frappe dans le système phonétique basque, c'est sa richesse. Les occlusives sont assez nombreuses, et il y a une grande variété de spirantes. Parmi celles-ci, on peut citer deux spirantes prépalatales. On signale aussi, sporadiquement, une cérébrale *ḍ* (entre *d* et *r*). Les phonèmes dits mouillés y sont fréquents. Le système des voyelles est, lui aussi, varié. Les voyelles nasales ne se rencontrent que dans un ou deux dialectes et le son de l'*ü* français que dans quelques localités voisines du domaine gascon et béarnais. Les diphtongues sont nombreuses, mais sont encore mal étudiées parce que l'étude de la prononciation basque a été faite pour une trop large part d'après des documents écrits.

L'accent basque est mal connu. Bornons-nous à dire ici que le souletin est le seul dialecte où certaines syllabes (généralement les pénultièmes) sont nettement plus intenses que les autres. Le chant de la phrase est assez différent suivant les dialectes et souvent même suivant les localités.

Les géminations de consonnes sont rares en basque, surtout à date ancienne.

Voici un spécimen qui donnera quelque idée de l'aspect phonétique du basque. Nous l'extrayons du *Nouveau Testament de Liçarrague* (1571) : c'est le début bien connu de la parabole de l'enfant prodigue (Luc XV, 11 et 12). (Nous en modernisons l'orthographe en nous efforçant de la conformer à la notation phonétique générale) :

11 *Halaber erran sesan, Gison batek situen bi seme :* [s = spirante prépalatale].

12 *Eta betarik gastenak erran sieson aitari, Aita, indak onhasunetik niri beltsen saitadan partea. Eta parti sietsen onak.*

Le système morphologique du basque est simple et varie seulement dans le détail suivant les dialectes. Au point de vue des procédés grammaticaux, c'est, avant tout, une langue à suffixation : le nombre des suffixes dont un mot peut être affecté peut aller

jusqu'à cinq ou six : exemple (que nous avons entendu de la bouche d'un enfant de quatre ans) : *ponetekilakoaekin* [*ponet-*(*e*)*kila-ko-a*(*re*)*-kin* « avec celui qui a le béret » (mot à mot : béret-avec-de-le-avec). Au point de vue des catégories grammaticales, voici quelques traits essentiels. Bien que le verbe puisse être affecté des mêmes suffixes que le nom — au moins dans la plupart des cas — la distinction du nom et du verbe est très marquée en basque. Très rarement un nom peut, avec un auxiliaire, constituer un verbe, mais l'auxiliaire est indispensable pour que le mot en question ait le sens verbal.

Le sujet de l'intransitif n'a aucune caractéristique, et c'est ce qui le fait reconnaître, car le *-k* qui se post-pose au sujet du verbe transitif n'est qu'un instrumental (l'homme bat l'enfant se dit en basque : par l'homme est battu l'enfant). Il suit de là que le complément direct du verbe n'existe pas. Quant au complément du nom, il le précède, tantôt avec l'un des suffixes qui indiquent le génitif, tantôt, quand le sens est suffisamment clair, sans caractéristique.

Il n'y a qu'une seule déclinaison en basque. Substantif, adjectif, formes verbales, pronom, quelques adverbes même sont susceptibles de se décliner, ceux-ci au moins partiellement. Le genre n'existe pas, et il n'est indiqué qu'en un seul cas dans une partie du verbe, pour indiquer le sexe d'une personne que l'on tutoie : du*k*, tu l'as [homme] ! du*n*, tu l'as [femme] !

Le verbe, si on le prend dans l'ensemble des dialectes et chez les auteurs, comporte une multitude de modes et de temps, mais dans la bouche des ruraux, il est à ce point de vue moins compliqué. De même pour la multiplicité des formes verbales, que les grammaires et les auteurs se plaisent à énumérer et quelquefois à forger. Normalement, le verbe est périphrastique (je suis en marche = je marche), mais il y a encore quelques verbes forts qui représentent l'ancienne conjugaison euskarienne : ils ne connaissent la périphrase qu'aux temps dits composés.

La syntaxe est très complexe et elle constitue la plus grande difficulté pour les étrangers qui veulent acquérir la pratique de

la langue basque, bien qu'elle ait dans une faible mesure subi l'influence des idiomes romans voisins. Pour en donner une idée, voici une phrase où l'ordre des éléments est presque toujours exactement l'inverse de ce qu'il est en français. Si je veux dire : « le livre que j'ai donné à l'enfant est très beau », voici comment je m'exprimerai (dial. guipuzcoan) : *Aurrari eman diodan liburua tšit ederra da*. L'analyse en est facile :

aurr-a-(r)i	*eman*	*diod-(a)n*	
enfant le à	donné	je l'ai de [c.-à-d. que]	
liburu-a	*tšit*	*ederr-a*	*da*
livre le	très	beau le	il est

Le vocabulaire basque a été trouvé fort riche par les uns, très pauvre par les autres. Il nous semble que la vérité se trouve entre ces deux extrêmes. Il est incontestable que le basque possède un nombre considérable de mots romans, et quelques vocables latins, celtiques, germaniques, arabes, etc., et que, fort souvent, lorsque l'on a à traduire un texte littéraire ou scientifique du français ou de l'espagnol, les mots manquent pour les rendre en basque. Mais, cela étant, il ne faut pas oublier qu'une foule de mots euskariens se sont perdus et que quotidiennement il s'en perd, en très grand nombre. Bien des expressions qu'on lit chez les auteurs anciens, ne sont plus comprises aujourd'hui, et il est d'observation courante que les vieillards ont un lexique plus riche que les jeunes. Par ailleurs, le basque possède à un degré éminent la faculté de former des mots nouveaux, et c'est par dizaines que l'on compte les suffixes de dérivation ; la composition est riche aussi. Et il est de fait que beaucoup d'auteurs ont forgé une quantité notable — quelquefois même trop — de mots nouveaux, qui sont parfaitement compris du vulgaire. Il faut cependant noter que si le peuple les entend, il n'en fait qu'un usage assez peu fréquent ; le prestige de l'espagnol et du français est tel en effet, que entre deux expressions, l'une purement basque (ou le paraissant), qu'elle soit ancienne ou forgée, et l'autre romanisante, le paysan basque a, le plus souvent, une tendance marquée à se servir de la seconde.

Les mots basques, en général, expriment des notions concrètes, mais ce n'est pas à dire, ainsi que des observateurs sachant médiocrement l'*euskera* l'ont soutenu avec véhémence, que le basque soit impropre à exprimer l'abstrait. Et il se peut que la sémantique basque réserve, à ce sujet, des surprises.

Georges LACOMBE.

BIBLIOGRAPHIE

On a écrit des centaines de livres, brochures et articles sur la langue basque, mais l'immense majorité d'entre eux ne peuvent être consultés qu'avec critique. Ici, nous sommes obligés d'énumérer seulement quelques ouvrages essentiels.

Parmi les bibliographies, la moins incomplète est l'*Essai d'une bibliographie de la langue basque*, de M. Julien Vinson. Paris, I, 1891 ; II, 1898. (Un 3e volume supplémentaire, annoncé, n'a pas paru.)

Les revues basques sont nombreuses : la plus sérieuse est la *Revue internationale des Études basques*, qui paraît depuis 1907.

Les deux parties de la grammaire basque qu'on a le plus étudiées sont la phonétique et le verbe. Pour la phonétique, l'ouvrage le plus recommandable est celui de M. Gavel : *Éléments de phonétique basque* (Paris et Biarritz, 1921). Pour le verbe, on trouve le plus de faits dans Bonaparte : *Le verbe basque*... (Londres, 1864 et 1869), avec ses deux compléments : *Études sur les dialectes d'Aezcoa, de Salazar et de Roncal*, Londres, 1872, et *The simple Tenses in modern basque and old basque*, Londres, 1884 [ouvrages malheureusement inachevés], et les théories les plus scientifiques dans H. Schuchardt : *Baskische Studien*. I : *Ueber die Entstehung der Bezugsformen des baskischen Zeitworts* (Vienne, 1893).

Pour le vocabulaire, on trouvera le plus de mots et d'exemples dans le *Diccionario vasco-español-frances* de M. de Azkue, 2 vol. (Paris et Bilbao, 1905-06).

Enfin, la meilleure documentation sur l'ibérien se trouve dans les *Monumenta linguae ibericae* de E. Hübner (Berlin, 1893), et dans *Die iberische Deklination* de M. Schuchardt (Vienne, 1907).

LANGUES CAUCASIQUES SEPTENTRIONALES

Généralités

On applique le nom de « langues caucasiques » à toutes les langues du Caucase qui ne sont ni indo-européennes, ni sémitiques, ni « touraniennes ». Ces langues forment deux grandes familles : les langues caucasiques septentrionales et les langues caucasiques méridionales. On suppose souvent une parenté entre ces familles, mais, comme cette parenté est loin d'être évidente, et comme elle n'a encore jamais été démontrée d'une manière vraiment scientifique, on fera mieux de n'employer le terme « langues caucasiques » que dans un sens géographique et de ne pas réunir les deux familles en question en un tableau généalogique.

Les langues caucasiques septentrionales sont parlées par un nombre d'individus qui doit être compris entre 1 et 2 millions. Elles peuvent être subdivisées en deux familles prochement apparentées l'une à l'autre : les langues *tchétchénolesghiennes*, à l'Est et les langues *abasgokerkètes*, à l'Ouest (voir la planche 8).

Toutes ces langues se distinguent au point de vue phonétique par la richesse de leur consonantisme : on notera surtout la multitude des consonnes labialisées, ainsi que les spirantes et les affriquées latérales (λ, *l̤*, *tl̤*) [1], inconnues à toutes les autres langues de l'Asie Occidentale.

En matière de morphologie, toutes les langues caucasiques

1. Les consonnes définies ici ne figurent pas dans le tableau général de transcription, non plus que celles définies page 336. Dans le présent chapitre *q* sert à noter une occlusive arrière-vélaire sourde (non emphatique) ; une apostrophe (') après une consonne indique l'occlusion glottale.

septentrionales s'accordent à distinguer le *casus agens* (sujet du verbe transitif), du *casus patiens* (sujet du verbe intransitif ou complément direct du verbe transitif) [1]. On notera encore que dans toutes ces langues les véritables racines verbales ne consistent qu'en une seule consonne : les exceptions à cette règle générale s'expliquent par la fusion d'anciennes racines « monolitères » avec certains préfixes, dont la fonction morphologique primitive était oubliée.

Outre ces traits communs de la phonétique et de la morphologie les langues tchétchénolesghiennes et abasgokerkètes possèdent aussi beaucoup d'éléments communs dans leurs vocabulaires.

Les langues tchétchénolesghiennes ou caucasiques-orientales.

Situation et caractéristiques.

Ces langues occupent le territoire contigu à la mer Caspienne et limité par le Terek au Nord et le Berbere-tchai au Sud.

Le consonantisme de ces langues est d'une richesse prodigieuse (l'*awar*, par ex., possède 43 consonnes), mais les rencontres de consonnes y sont évitées, à l'exception des combinaisons les plus simples (liquide + consonne, etc.). Le vocalisme, au contraire, est assez pauvre. Toutes les langues de cette famille ont un accent expiratoire, qui dans la plupart des cas est si faible qu'on le remarque à peine.

Le trait le plus original de la morphologie des langues tchétchénolesghiennes consiste en ce que tous les substantifs y sont répartis entre plusieurs « classes » ou genres grammaticaux, dont le nombre s'élève parfois jusqu'à six. A part le « masculin » et le « féminin », ces genres ne se rattachent à aucune catégorie sémantique déterminée, et si dans certaines langues on remarque la tendance à rationaliser le système des genres au point de

1. Voir à ce sujet l'article très instructif de M. Schuchardt, *Sitzungsberichte der k. k. Wiener Akademie, phil. hist. Kl.* 133, 1.

vue sémantique, la comparaison de ces langues entre elles prouve que ce n'est qu'une innovation secondaire. Chaque genre est caractérisé par une consonne *w*, *v*, *y*, *d*, *r*, *l*, *b* ou par deux de ces consonnes (l'une pour le singulier, l'autre pour le pluriel). Ces consonnes, préfixées, infixées ou suffixées à certains mots (adjectifs, pronoms démonstratifs, verbes, substantifs, noms de nombre, substantifs au cas génitif ou locatif), établissent l'accord en genre et en nombre de ces mots avec les substantifs auxquels ils se rapportent. Ainsi en awar où *w* caractérise le masculin et *y* le féminin, on dit : *λ'abaugo tš'i* « trois hommes », mais *λ'abaigo ttšužu* « trois femmes », *ḫerau tš'i* « vieil homme », *ḫerai ttšužu* « vieille femme » ; les verbes s'accordent avec leur *patiens* ; *tš'i vix'izavize* « montrer un homme », *ttšužu yix'izayize* « montrer une femme » ; les « locatifs » (latifs, ablatifs) avec l'objet, dont ils marquent la position dans l'espace *tš'i vugo roq'ou* « l'homme est dans la chambre », *ttšužu yigo roq'oi* « la femme est dans la chambre » [1].

La flexion nominale et verbale s'opère par des suffixes (ou désinences), ajoutés à des thèmes différents.

Le verbe peut avoir deux thèmes, celui du « duratif » et celui du « ponctuel », dont la formation varie d'un verbe à l'autre : ḥurq'ili (*d-*, *ƀ-*) *udžis'* – (*d-*, *ƀ-*) *irdžis'* (ponct.) « boire », mais (*w-*, *r-*, *d-*, *ƀ-*) *irq' wis'* (durat.) – (*w-*, *r-*, *d-*, *ƀ-*) *arq'wis'* (ponct.) « faire » etc.

Les substantifs peuvent avoir quatre thèmes, — celui du « patiens » singulier, du « patiens » pluriel, des autres cas du singulier et des autres cas du pluriel [2], — et la formation de ces thèmes varie d'un substantif à l'autre : lak' pat. sing. *barts'* « loup » gén. sing. *burts'i-l*, pat. plur. *barts'ru*, gén. plur.

1. Voir à ce sujet A. Dirr, *Ueber die Klassen (Geschlechter) in den Kaukasischen Sprachen* dans les *Archives Internationales d'Ethnographie*, XVIII (Leide, 1908), pp. 125 et ss.

2. Il arrive parfois que certains substantifs présentent encore des thèmes supplémentaires pour certains cas du singulier. Ainsi en awar *rox'* « bois », est au génitif *rox'i-l* et à l'ablatif *rox'o-sa*, tandis que *ḫor* « lac » qui a le génitif *ḫori-* présente à l'ablatif *ḫorini-sa*.

burts'irdi-l, mais d'autre part pat. sing. *lag* « esclave » gén. sing. *lagnal*, pat. plur. *lagart'*, gén. plur. *lagart'undal*.

On voit que, pour les verbes comme pour les noms, les différents thèmes ne sont liés entre eux que par l'identité des consonnes radicales.

Les pronoms personnels ont une flexion encore plus « irrégulière » : hurq''li pat. *nu* « moi » — agens *dali* ; awar pat. *mun* « toi » — agens *dutsa*. Il n'y a que les pronoms démonstratifs, les adjectifs et les noms de nombre qui présentent une flexion plus ou moins uniforme, sans alternance de thèmes.

Le système de la déclinaison tchétchénolesghienne est richement détaillé. Outre les cas à valeur relativement abstraite (« casus agens », « casus patiens », génitif, datif, comitatif, comparatif, etc.), on y trouve un très grand nombre de cas « locaux » plus concrets. La plupart des nuances qui sont traduites dans les langues indo-européennes par des prépositions sont exprimées dans les langues tchétchénolesghiennes par des formes casuelles spéciales : en awar on dit (*bos'ize*) *got'otlan* « (prendre) de dessous l'arbre » ou (*ine*) *rox'iq'ehun* « (aller) dans la direction du bois ». Le nombre des cas devient parfois énorme : l'awar par exemple en possède trente.

Les langues tchétchénolesghiennes ne connaissent point la distinction fondamentale entre les formes personnelles et les formes impersonnelles du verbe. Toutes les formes de certains verbes s'accordent en genre et en nombre avec leur patiens. L'accord en « personne » est inconnu à la plupart des langues tchétchénolesghiennes ; de véritables désinences personnelles n'existent que dans deux langues (le lak' et le dargwa) ; deux autres (l'udi et le t'abasaran) les remplacent par des formes enclitiques des pronoms personnels correspondants.

Divisions.

Les langues tchétchénolesghiennes peuvent être divisées en huit groupes [1] :

1. Les véritables *nomina ethnica* sont fort rares dans le Caucase Oriental.

I. — Le groupe *tchétchène* embrasse trois branches :

a) le *tchétchène* (*naḫtš'uin muott'*) proprement dit avec ses dialectes médiocrement étudiés (*galgaï, a'uḫ, itškeri,* etc.) est répandu dans les districts (*okrugi*) de Grozny, Védène et Kassav-Yourt de la « Région du Terek » (*Terskaja oblastj*) ;

b) l'*inguš* dans le district de Nazran (dans les environs de Vladikavkaz) ;

c) le *t'uš* ou *bats'* dans quelques villages du bassin de l'Alazan supérieur dans le district de Telav (en Géorgie).

Tous ces idiomes se ressemblent de près. Tous les trois possèdent un système de 6 genres grammaticaux, dont 2 seulement (le masculin et le féminin) coïncident avec des catégories sémantiques. Chaque substantif présente pour différents cas différents thèmes dont la ormation varie selon le substantif; dans la conjugaison chaque verbe possède au moins deux thèmes, celui de l'aspect duratif et celui du ponctuel. En tchétchène proprement dit cette hypertrophie de différenciation morphologique est encore accrue par l'effet de certaines lois phonétiques (contractions, chutes de certaines consonnes intervocaliques, influence du vocalisme de la seconde syllabe sur celui de la première) : ainsi *t'ur* « sabre » a le génitif *t'üiriñ*, l'instrumental *t'üiruö*, le pluriel *t'arraš*, tandis que *sai* « cerf » présente le génitif *sēñ*, l'instrumental *sē* et le pluriel *siēš*. Grâce à ces circonstances, presque chaque nom et une grande partie des verbes sont « irréguliers », ce qui — en combinaison avec le système de six genres grammaticaux — fait du tchétchène la langue la plus difficile de tous les idiomes caucasiques. C'est pourquoi, peut-être, cette langue ne tend guère à se propager au delà de ses limites primitives, bien que les Tchétchènes soient le

La plupart des habitants du Daghestan se désignent eux-mêmes comme « montagnards » ou « habitants de tel ou tel village ». Cette circonstance est très embarrassante pour le linguiste qui doit établir une nomenclature scientifique, et on se voit souvent forcé de recourir à des procédés plus ou moins artificiels. Voir sur cette question A. Dirr, *Die heutigen Namen der Kaukasischen Völker*, *Petermann's Mitteilungen aus Justus Perthes geographischer Anstalt*, 1908, IX, pp. 204 et suiv.

peuple le plus nombreux du Caucase du Nord (300.000). — Par son vocabulaire le groupe tchétchène occupe une position isolée : le nombre de mots tchétchènes qui trouvent des équivalents exacts dans les autres langues tchétchénolesghiennes est relativement petit.

II. — Le groupe *awaroandi* embrasse trois branches :

a) l'*awar* ou *ma'arul matts* « (langue des montagnes ») avec plusieurs parlers locaux est répandu comme langue natale des indigènes dans les districts de Gounib et d'Avar (dans le Daghestan) ainsi que dans le Nord du district de Zakataly [1] ;

b) les langues de la vallée du bas Koïssou d'Andi — le *'andi* (ou *q'uannu*), le *bo̊λiq'*, le *godoberi*, le *tšamalal*, le *k'arat'a*, le *åλuaλ*, le *qüanada* (ou *bagulal*) et le *t'indi* — dans le district d'Andi ;

c) les langues du Koïssou d'Andi supérieur (du « Didoet'i ») — le *dido*, le *ḫwarši*, le *qaputši* — dans le même district [2].

Toutes les langues awaroandi, liées entre elles par une parenté intime, manifestent une tendance à simplifier leur structure morphologique, tout en conservant leur originalité. Le nombre des genres grammaticaux est réduit à trois (masculin, féminin, neutre) dans tous ces idiomes à l'exception du 'andi qui en possède quatre. Dans la conjugaison, la distinction des aspects duratif et ponctuel n'existe plus. Dans la déclinaison on remarque une tendance à éliminer la multiplicité des paradigmes en remplaçant la flexion nominale par la flexion pronominale, et en awar, par exemple, presque chaque substantif peut être décliné d'après le modèle des pronoms démonstratifs.

Au point de vue phonétique les langues de ce groupe se distinguent par des affriquées vélaires (*k'ḫ* et *k'ḫ*) et par la multiplicité des consonnes latérales (λ, λ', λλ, *tl*). — L'awar est la langue la plus importante de ce groupe : elle sert de « lingua franca » non seulement aux indigènes du district d'Andi, dont les

1. Voir A. Dirr, *Die Bevölkerung des Sakataler Kreises*, *Petermanns Mitteilungen*, 1915, VIII (avec carte), pp. 309 et ss.

2. Pour la répartition des langues du bassin du Koïssou d'Andi, v. A. Dirr, *Sprachenkarte der Gebiete am Mittellauf des Andischen Koïssu (Daghestan)*, *Peterm. Mitt.*, 1907, X, pp. 234 et ss.

langues natales lui ressemblent de près, mais aussi aux habitants des districts de Kazikoumoukh et de Dargi, où elle entre en concurrence avec le *koumyk* (famille turque), la langue de commerce de tout le Nord-Est du Caucase.

III. — Le *lak'* ou *kazikoumoukh* est parlé avec des variations dialectales peu importantes dans tout le district de Kazikoumoukh (Daghestan Central).

Au point de vue phonétique, cette langue est caractérisée par la pauvreté de son vocalisme et par sa prédilection pour les consonnes géminées.

Dans la déclinaison on signale la confusion du *casus agens* avec le génitif et la distinction sévèrement maintenue des 4 thèmes nominaux (2 pour le singulier et 2 pour le pluriel), dont les types de formation sont très multiples. Le verbe possède des désinences personnelles. Les substantifs se répartissent entre quatre genres grammaticaux.

IV. — Le groupe *dargwa* embrasse les idiomes du district de Dargi et du Nord-Ouest du district Kaïtagotabassaran (Daghestan Oriental). Les villages de cette région forment des sortes de confédérations appelées *dargwa*. Les noms de ces dargwa (*Aq'uša-dargwa*, *Varq'un-d.*, *Qăƀa-d.*, etc.) ne semblent pas toujours indiquer des unités linguistiques. Les « Dargwines » parlent des idiomes très semblables, dont seul le *qubatši* (ou *'ärƀuk'*) semble occuper une place à part. Le seul idiome dargwa bien étudié jusqu'à présent est le *ḥurq'ili*. De toutes les langues tchétchénolesghiennes le *ḥurq'ili* semble avoir le plus fidèlement conservé la multiplicité des formations des thèmes d'aspect duratif et ponctuel. Tout comme le lak', le ḥurq'ili possède des désinences personnelles mais, en partie du moins, elles sont issues de pronoms enclitiques. La déclinaison est simplifiée : les types de formation du pluriel sont encore assez multiples, mais pour le singulier tous les substantifs ne présentent qu'un seul paradigme avec des variations d'origine purement phonétique.

V. — Le groupe *samourien* [1] embrasse quatre branches :

1. Pour la répartition géographique des langues de ce groupe et des trois

a) les langues du Nord-Est, — le *k'üri* (avec ses deux dialectes, *k'üri* proprement dit et *aḫti*) dans la partie orientale du district de Samour, l'Est du district de Kuri et le Nord du district de Kouba (gouvernement de Bakou), l'*agul*, occupant le courant supérieur du Gulghary-tchaï et de son affluent, le Kourakh-tchaï, et le *t'abasaran* dans le bassin de Rubastchaï ;

b) les langues du Sud-Est — le *buduḫ* et le *dẑek* (*ḫaputli* ou *qriẑ*), parlés au pied du Chahdagh ainsi que dans quelques villages au bord de la mer Caspienne ;

c) le *rut'ul* (ou *meḫ'əd*), parlé dans le district de Samour à l'Ouest du k'üri ;

d) le *ts'aḫur* — dans le même district, à l'Ouest du *rut'ul*, ainsi que dans quelques villages du district de Zakataly [1].

Toutes les langues samouriennes tendent à simplifier leur structure morphologique et phonétique.

A l'Ouest cette tendance est encore faible : le consonantisme du rut'ul et du ts'ahur est assez simple, mais la morphologie de ces langues conserve encore son aspect tchétchénolesghien (quatre genres grammaticaux, multiplicité des paradigmes, etc.).

Les langues du Nord-Est présentent une structure morphologique toute autre. Le t'abasaran ne possède que deux genres « humain » et « non-humain », le k'üri et l'agul ne distinguent plus du tout les genres. Le k'üri s'efforce de combattre un autre grand inconvénient de la grammaire caucasique-orientale, la multiplicité des paradigmes de la déclinaison nominale. Il n'a conservé (ainsi que le t'abasaran) qu'un seul type de formation du pluriel pour tous les substantifs. Le singulier comporte encore sept types différents de la formation du thème des cas obliques, mais on remarque une tendance à rationaliser la répartition des substantifs entre ces types de paradigmes ; ainsi, les substantifs polysyllabiques forment leur « agens »

groupes suivants (VI, VII, VIII), voir A. Dirr, *Karte der kürinischen Sprachgruppe*, *Peterm. Mitt.*, 1909, II, pp. 138 et ss.

1. Voir A. Dirr, *Die Bevölkerung des Sakataler Kreises (mit Karte)*, *Peterm. Mitt.*, 1915, VIII, pp. 309 ss.

en *-di* (pat. *dzanáwur* « loup » — ag. *dzanáwurdi*), les monosyllabiques désignant des matières les forment en *-ádi* (pat. *rug* « argile » — ag. *rukwádi*), ceux qui désignent des animaux les forment en *-rá* (pat. *lam* « âne » — ag. *lamrá*), etc. Les exceptions sont encore nombreuses, mais la tendance générale est bien marquée et semble continuer encore à exercer son action.

Au point de vue phonétique, le k'üri est caractérisé par son accent expiratoire très fort, qui repose le plus souvent sur la seconde syllabe du mot et exerce une grande influence non seulement sur les voyelles, mais aussi sur les consonnes.

Les langues samouriennes sont en train de disparaître, cédant le terrain au turc-azéri (azerbeïdjan). Seul le k'üri tient encore tête à cette langue, qui est la « lingua franca » de tout le Caucase méridional.

VI. — L'*artši* n'est parlé que dans le village Artchi (800 habitants) sur la frontière sud-ouest du district de Kazikoumoukh au bord de la rivière Khotar (affluent du Koïssou de Kazikoumoukh). Malgré leur nombre fort restreint, les habitants de ce village ont réussi, grâce à leur situation isolée et à une endogamie sévère, à conserver leur langue maternelle. L'influence de l'awar d'une part et du lak' de l'autre se fait sentir, il est vrai; l'artši reste néanmoins un idiome indépendant ne faisant partie d'aucun autre groupe de langues tchétchénolesghiennes. Très archaïque sous beaucoup de rapports, cette langue est d'une grande importance pour la grammaire comparée des langues tchétchénolesghiennes.

VII. — L'*udi* n'est parlé que dans deux villages (Niche et Vartachène) du district de Noukha (gouv. Elisavetpol). Les « oudines » sont chrétiens (arméno-grégoriens). Le vocabulaire de leur langue — à part les éléments arméniens assez nombreux — ressemble beaucoup à celui des langues samouriennes. Mais la phonétique et la morphologie ne permettent pas de classer l'*udi* parmi ces langues.

L'*udi* ne distingue plus les genres grammaticaux et les aspects duratif et ponctuel. Le verbe possède des désinences personnelles qui ne sont que des formes enclitiques des pronoms correspondants. Dans la déclinaison on notera l'existence d'un accusatif différent du *patiens* et de l'*agens*.

VIII. — Le *ḫinalug*, parlé dans un seul village du même nom au pied de Chahdagh n'est pas encore suffisamment étudié. Tout ce qu'on connait de cette langue est une petite liste de mots, qui permet d'affirmer que c'est une langue indépendante, ne faisant partie d'aucun autre groupe des langues tchetchénolesghiennes.

Langues abasgokerkètes ou caucasiques occidentales.

Caractéristiques.

Les langues abasgokerkètes (ou caucasiques occidentales) sont caractérisées au point de vue phonétique par la multiplicité des groupes de consonnes. Ce trait distingue ces langues de leurs sœurs du Caucase oriental, qui, on l'a vu, n'admettent que les groupes de consonnes les plus simples. On notera dans les langues abasgokerkètes des chuintantes et des demi-chuintantes labialisées ($\check{s}^w$, $ṣ^w$, etc.), presque inconnues aux langues tchétchénolesghiennes (à l'exception du t'abasaran). L'abkhaz et l'oubykh possèdent des vibrantes labiales (sonore ꝧ et sourde ꞇ), pareilles aux sons que produisent les cochers pour arrêter les chevaux : abkh. *ꝧį* « champ », oub ykhꞇ*uk* « cou ».

L'abkhaz n'a pas de déclinaison ; l'oubykh et l'adyghé ne distinguent que deux cas, dont l'un sert de « patiens » et l'autre à la fois d'« agens » et d'« obliquus » (génitif, datif, locatif, ablatif) : ces faits sont en contraste profond avec l'exubérance de la déclinaison tchétchénolesghienne. Toutes les langues abasgokerkètes possèdent un article défini, inconnu aux langues caucasiques orientales. La notion des genres grammaticaux est étrangère aux langues abasgokerkètes, mais il y a une série de faits qui prouvent qu'il n'en a pas toujours été ainsi : l'abkhaz, par exemple, a deux préfixes personnels de la troisième personne, l'un pour les « êtres humains », l'autre pour les « non-humains », etc.

Un des traits les plus originaux de la morphologie abasgokerkète sont les « préfixes personnels ». Ces préfixes, qui renferment toujours la consonne radicale du pronom personnel correspondant, sont préposés aux substantifs pour indiquer la possession,

et aux formes verbales pour indiquer l'agens, le patiens et le complément indirect. Ainsi, en circassien, où *sə* veut dire « moi » et *wə* — « toi », on dit : *ṣ̂ʷüiz̧* « femme », *si-ṣ̂ʷüiz̧* « ma femme », *wi-ṣ̂ʷüiz̧* « ta femme », *we-s-t'yn* « je te donnerai », etc.

Un autre trait caractéristique est la prédilection de ces langues pour les mots composés, qui expriment parfois les notions les plus simples : p. ex. qabardi *z̧a-k'e* « barbe » (= « queue de la bouche »), *na-p'e* « visage » (= « œil et nez »), *ne-p's* « larme » (= « eau de l'œil »), etc.

Divisions.

La famille abasgokerkète comprend trois groupes linguistiques.

I. — L'*adyghé* (*adyge-yabz̧e*), qui se partage en deux branches :

a) le *qabardi* (*qeberdei-yabz̧e*) ou « haut adyghé », principalement dans le district de Naltchik (rég. du Terek) ;

b) le « bas adyghé », (*k'yax'ə-abz̧ə*) ou circassien (tcherkesse), parlé jadis dans toute la steppe au Sud du Kouban, à partir de la Tiberda jusqu'à la mer Noire, et sur le bord de la mer entre l'embouchure du Kouban et la rivière Chakhé, et comprenant deux groupes de parlers, — celui de la steppe (dialectes *abz̧ax'ə*, *bž̧eduḫʷ*, etc.), et celui du littoral (dial. *ṣyap'sʐg*, *net'ḫu-qwadž̧ə*, *ḫak'utšʷ*) ; la plupart des Circassiens ayant émigré en Turquie après la conquête du Caucase par les Russes, il ne reste maintenant au Caucase qu'une vingtaine de villages circassiens dans la région du Kouban (distr. Batalpachinsk, Laba, Maïkop) et dans le district de Touapsé (gouv. de la Mer Noire).

Les deux branches de l'adyghé sont très proches l'une de l'autre, et l'« adyghé commun » se laisse facilement reconstruire. La phonétique de l'adyghé est moins bizarre que celle des autres langues abasgokerkètes. En morphologie on notera l'article postposé *-r* (qab. *z̧ak'e* « barbe », *z̧ak'er* « la barbe ») et les verbes composés de deux racines verbales (qab. *nesịn* « arriver », *ž̧en* « courir » — *než̧esịn* « accourir, arriver en courant », etc.).

II. — L'*oubykh* (*a-ʐoḫ*) avec ses deux dialectes (*saṣe* et *vardane*)

occupait jadis le bord de la Mer Noire entre les rivières Chakhe et Chatché. Après la conquête russe tous les oubykh émigrèrent en Turquie. Leur langue, très peu connue, semble être un intermédiaire entre le circassien et l'abkhaz. Elle a subi une très forte influence circassienne d'abord, turque ensuite. A l'heure qu'il est, la plupart des oubykh en Asie Mineure ont complètement oublié leur langue et ne parlent que le turc.

III. — L'*abkhaz*, comportant plusieurs dialectes, dont les rapports mutuels ne sont pas encore suffisamment étudiés, est parlé sur le bord de la Mer Noire entre les rivières Chatché et Ingur et dans le hinterland de cette région, ainsi que dans quelques villages épars sur les bords du Kouban et du Kouma (district Batalpachinsk).

Au point de vue phonétique c'est bien la langue la plus difficile et la moins harmonieuse de tout le Caucase. Sa morphologie est caractérisée par l'absence de déclinaison et par l'incorporation de certains substantifs dans le corps des verbes : *tš'y* « bouche », — racine verbale *ts'* « mettre » — *il-tš'a-s-ts'-ueit'* « je (*-s-*) lui (*-l-*) le (*-i-*) mets dans la bouche » (*-ueit'* = désinence du présent).

Écriture et littérature

Les indigènes du Caucase Septentrional étant musulmans pour la plupart, quelques-uns d'entre eux ont essayé d'écrire leurs langues natales en se servant de l'alphabet arabe. Mais, ce dernier n'étant pas approprié à la richesse du consonantisme caucasique, toutes ces tentatives durent échouer. Il en fut de même pour les tentatives de quelques tribus chrétiennes (les Abkhaz, les Udi, les T'uš) qui essayèrent d'adapter à leurs langues les alphabets géorgiens ou arméniens. Le gouvernement russe essaya vainement à plusieurs reprises de donner, surtout aux montagnards du Daghestan, un alphabet phonétique, reposant sur l'alphabet russe, avec certaines modifications : on se heurta à la résistance du clergé musulman, qui soupçonnait toujours des tentatives de russification ou de christianisation. C'est ainsi qu'aucune des

langues caucasiques septentrionales ne possède ni écriture nationale ni littérature écrite.

Histoire de l'étude des langues caucasiques septentrionales.

Il n'existe pas jusqu'à présent de grammaire comparée des langues caucasiques septentrionales qui soit à la hauteur des exigences de la méthode scientifique [1]. Par conséquent, toutes les spéculations étymologiques et en partie glottogoniques sur ces langues entreprises par certains savants (Winkler, Bork, Gleye, Marr, Trombetti, etc.) sont prématurées et manquent de base scientifique. On aura soin de s'en méfier.

Les ouvrages purement descriptifs ne sont pas tous de valeur égale. Ainsi les travaux de Klaproth (*Reise in dem Kaukasus*, 1812-1814, *Tableau historique...du Caucase*, 1827) et de Erkert (*Die Sprachen des Kaukasischen Stammes*, Wien. 1895) sont pleins de fautes et ne peuvent qu'induire en erreur celui qui voudra les consulter.

Pour les langues tchétchénolesghiennes on ne peut recommander que les grammaires du baron P. Uslar, A. Schiefner et A. Dirr. — Les cinq grammaires du baron Uslar parurent d'abord traduites en allemand par A. Schiefner. Ce n'est qu'après la mort du baron Uslar que le texte original de ses ouvrages fut publié à Tiflis par la Direction de l'Instruction Publique du Caucase sous le titre *Etnografija Kavkaza*, vol. II, *Tšetšenskij jazyk*, 1888 (= A. Schiefner, *Bericht über Baron P. K. Uslars tschetschenzische Studien*, Saint-Pétersbourg, 1867); vol. III, *Avarskij jazyk*, 1889 (= A. Schiefner, *Bericht... awarische Studien*, 1872); vol. IV, *Lakskij jazyk*, 1890 (= A. Schiefner, *Bericht... kasikumükische Studien*, 1866); vol. V, *Chjurkilinskij jazyk*, 1892 (=A. Schiefner, *Bericht.., hürkanische Studien*, 1871); vol. VI, *Kjurinskij jazyk*, 1896 (= A. Schiefner, *Bericht... kürinische Studien*, 1873). Toutes ces grammaires sont suivies de textes et de dictionnaires. Leur

1. Voir *Bulletin de la Société de Linguistique*, t. XXIII, pp. 184 et ss.

seul côté faible est la description et la transcription des phonèmes caucasiques. Ne possédant pas l'érudition linguistique nécessaire, le baron Uslar n'a pas su se servir de la transcription phonétique proposée par l'Académie de Saint-Pétersbourg : ainsi, par exemple, il transcrit les explosives aspirées *t*ʻ, *p*ʻ par des lettres qui doivent indiquer des inaspirées, et les explosives à occlusion glottale *t*', *p*', *k*' par des lettres indiquant des aspirées, etc. Ce n'est qu'après une analyse critique très attentive qu'on parvient à déchiffrer cette transcription et à déterminer la valeur véritable des phonèmes qu'elle représente [1]. Les remarques sur la phonétique du k'üri et du dargwi qui accompagnent les textes k'üri du prince Lionidze et de B. Soultanof et les textes « tsoudakars » (= dargwi) de B. Dalgat (*Sbornik materialov dlja opisanija mèstnostej i plemen Kavkaza*, v. XIV) viennent en aide à cette analyse.

Parmi les ouvrages de A. Schiefner, les *Awarische Texte* (St. Ptb., 1873) sont écrits dans la même transcription que les textes de la grammaire awar du baron Uslar. Le *Versuch über die Sprache der Uden* (1863) et le *Versuch über die Thusch-Sprache* (1856) présentent une transcription plus exacte, quoique dans ces deux ouvrages A. Schiefner se trouve sous l'influence de ses informeurs arméniens et géorgiens, qui ne sentent pas — on le sait — la différence entre *t* et *t*' ou entre *ts* et *ts*ʻ. C'est à cette influence aussi qu'il faut attribuer le fait que les dictionnaires qui accompagnent ces deux ouvrages contiennent une foule de mots arméniens et géorgiens, qui sans doute n'existent pas dans le langage familier des Udis et des Tʻuš.

Tous ces défauts n'existent pas dans les remarquables grammaires de M. A. Dirr : 1. — *Udinskaja grammatika* (*Sbornik materialov dlja opisanija mèstnostej i plemen Kavkaza*, XXXIII, 1904); 2. — *Grammatitšeskij otšerk tabasaranskago jazyka* (Ibid., XXXV, 1905) ; 3. — *Kratkij grammatitš. otšerk andijskago jazyka* (Ibid., XXXVI, 1906); 4. — *Aguljskij jyzyk* (Ibid., XXXVII, 1907); 5 — *Artšinsk. jazykij* (Ibid., XXXIX, 1908); 6 —

1. Cette analyse critique ayant été produite par l'auteur du présent chapitre, tous les mots caucasiques cités ci-dessus sont écrits non dans la transcription d'Uslar, mais dans une transcription plus exacte. Voir l'article cité à la note de la page précédente.

Materialy dlja izutšenija jazykov i narètšij ando-didojskoj gruppy (Ibid., XL, 1909) ; 7 — *Rutuljskij jazyk* (Ibid., XLII, 1912) ; 8 — *Cachurskij jazyk* (Ibid., XLIII, 1913). Toutes ces huit grammaires, suivies de textes et de dictionnaires, peuvent être recommandées sans aucune réserve. Les pp. 157-167 de la grammaire de l'*agul* (*Sbornik*, XXXVII, 1907) contiennent une liste de mots des langues du Daghestan méridional, que M. Dirr réunit sous le nom de « langues kurines » (langues samouriennes, udi, artši et hinalug) : pour le džek, le buduh et le hinalug c'est tout ce qu'on possède de positif jusqu'à présent.

Parmi les langues abasgokerkètes il n'y a que le qabardi qui se prête aux études scientifiques grâce aux nombreuses publications de MM. L. Lopatinski et P. Tambiev. La grammaire qabardi, rédigée par Lopatinski et publiée dans le *Sbornik materialov*, etc., XII (1891), laisse beaucoup à désirer au point de vue scientifique; les commentaires, qui accompagnent les textes qabardis publiés par Lopatinski et Tambiev dans différents volumes du même *Sbornik materialov*, sont pleins d'erreurs importantes. Mais les textes eux-mêmes sont bons et on peut en dire autant du dictionnaire russe-qabardi qui suit la grammaire de Lopatinski [1].

Pour le circassien on ne possède à vrai dire que le livre de M. Lulié, *Slovarj russko-tšerkesskij s kratkoj grammatikoj*, Odessa 1846, dont l'auteur s'est proposé la tâche inexécutable de rendre les phénomènes multiples du circassien par les lettres de l'alphabet russe sans recourir à des signes supplémentaires, et de traiter la morphologie très originale de cette langue dans les termes de la grammaire scolaire russe. Les textes soi-disant « circassiens », publiés par M. P. Tambiev (Qabardi de naissance) dans divers volumes du *Sbornik materialov*, ne donnent aucune idée de la langue circassienne : ils sont rédigés en partie en

1. Le « Kabard nyelv-táj » de M. Balint ne fait que reproduire en hongrois la grammaire de Lopatinski. Quant au « Dictionarium kabardico-hungarolatinum », du même auteur, c'est un livre dont on aura soin de ne pas faire usage : les erreurs de M. Lopatinski y sont encore renforcées par les théories de M. Balint, qui nie la parenté du hongrois avec les langues finno-ougriennes et réunit en un seul groupe le hongrois, le qabardi et... le japonais (!).

qabardi pur, en partie en une sorte de jargon circassien-qabardi, comparable au « français » des guides italiens ; les vrais Circassiens ne les comprennent pas ; l'auteur du présent chapitre a eu l'occasion de s'en convaincre personnellement [1].

Pour l'oubykh on ne possède jusqu'à présent que quelques notices du baron P. Uslar (*Etnografia Kavkaza*, I, 1887) et un court compte rendu de M. A. Dirr (imprimés dans les *Aufsätze zur Kultur- und Sprachgeschichte, Ernst Kuhn zum 70 Geburtstag 7. II. 1916 gewidmet*, pp. 413-419) ; il est permis d'espérer que cet éminent caucasologue, qui a étudié l'oubykh sur place (dans les environs d'Izmid en Asie Mineure), ne tardera pas à publier in-extenso ses études sur cette langue presque disparue.

L'abkhaz a été traité par le baron P. Uslar (*Etnogr. Kavk. I : Abchazskij jazyke*, 1887 = A. Schiefner, *Bericht über Baron P. Uslars abchasische Studien*, 1863), mais fort superficiellement. M. Marr a promis de publier une nouvelle grammaire de cette langue, et il faut espérer que les théories linguistiques un peu extravagantes de cet éminent caucasologue ne l'empêcheront pas d'en donner un tableau exact et objectif.

Prince N. TROUBETZKOY.

1. Toutes les données sur le circassien que renferme le présent chapitre sont tirées des travaux de son auteur consacrés à cette langue, travaux qui jusqu'à présent n'ont pu être publiés.

LANGUES CAUCASIQUES MÉRIDIONALES

Quatre langues parlées au Sud du Caucase forment un groupe défini, qu'on peut nommer *k'art'vélien*, d'après le nom indigène du principal idiome du groupe, le géorgien : le *géorgien* (en russe *gruzin*), le *mingrélien*, le *svane* et le *laze*. Ces quatre langues représentent une même langue commune. Toutefois le laze et surtout le svane sont bien différents du géorgien.

Les Grecs nommaient Ibères la population de la région où se parlent ces langues, et les Arméniens la nomment *Virk'*, ce qui est le même mot ; car le *v* arménien peut représenter un ancien *b* intervocalique, et *i* un ancien *ē*.

Le nombre des sujets parlant ces langues est d'environ un million et demi. Il s'agit sans doute de la survivance d'une population autrefois plus étendue. En particulier le laze, qui est parlé par un petit nombre de musulmans sur la partie sud-est de la côte de la mer Noire, de Batum à Platana, n'est évidemment que le débris d'un groupe qui a dû avoir plus d'importance.

Les Géorgiens sont les seuls à posséder une langue écrite. Il y a, depuis le xe siècle ap. J.-C. au moins, une littérature chrétienne, composée en une langue fixée. Il y a un alphabet propre au géorgien, formé comme l'alphabet arménien de l'alphabet grec, mais avec addition d'autant de signes qu'il est nécessaire pour donner à chacun des phonèmes du géorgien une notation propre.

Comme l'arménien, le géorgien a trois séries d'occlusives, une série sourde, soit *t*, une série sourde aspirée, soit *t'*, et une série sonore, soit *d*. Et ces trois séries existent non seulement pour les gutturales, les dentales et les labiales, mais aussi pour

les mi-occlusives, du type sifflant, comme *c* (*ts*), ou du type chuintant comme *č* (*tš*). De plus, le géorgien a une laryngale qui se note en transcription par *q*. Les voyelles sont les voyelles les plus usuelles *a*, *e*, *i*, *o*, *u*, sans indication de nuances. C'est le consonantisme qui donne à ces langues leur caractère phonétique propre.

Les caractéristiques morphologiques sont nombreuses et variées. L'élément radical peut être entouré à la fois de préfixes et de suffixes. Le verbe est distingué du nom par des affixes nets. On aura une idée du fonctionnement des affixes par la comparaison de *k'art'vel-i* « Géorgien » et du dérivé *sa-k'art'vel-o* « Géorgie ».

La forme même du nom sert à la fois de cas sujet et de complément direct. La seule forme casuelle qui s'oppose à cette forme principale du nom est celle d'un cas complément du nom caractérisé par *-i* (voir Reby, *M S L*, XVIII, p. 219 et suiv.).

Le verbe admet deux types, dont l'un appelle une forme de cas-sujet pourvue d'un suffixe, et semble n'être pas proprement transitif.

En ce qui concerne la structure générale, le caucasique méridional est intermédiaire entre le type flexionnel indo-européen et le type turc ou finno-ougrien à caractéristiques toutes autonomes. La flexion verbale est très complexe. Pour la forme du nom, les caractéristiques casuelles sont les mêmes au pluriel qu'au singulier, ainsi *mama* « père », génitif *mamis*(*a*) ; plur. *mamebi*, génitif *mamebis*(*a*), etc. Mais il y a deux types de formation du pluriel des noms.

A. MEILLET.

INDICATIONS BIBLIOGRAPHIQUES

Le petit exposé de Finck, *Die Haupttypen des Sprachbaus*, p. 132-146, permettra de pénétrer plus avant dans la structure du géorgien. La grammaire la plus recommandable est : A. Dirr, *Theoretisch-praktische Grammatik der georgischen Sprache* (Vienne, collection Hartleben). Le dictionnaire géorgien à consulter est celui de *Tchubinov*, où les mots géorgiens sont interprétés par le russe et le français.

LANGUES DRAVIDIENNES

La famille dravidienne.

Des trois groupes de langues actuellement parlées dans l'Inde, l'un, l'indo-aryen, une forme de l'indo-européen, est certainement d'origine étrangère ; l'autre, le muṇḍā, s'apparente à un groupe indochinois ; seul le dravidien, qui est confiné dans l'Inde, et presque entièrement dans l'Inde péninsulaire, a une existence indépendante, et toutes les tentatives faites jusqu'ici pour le rattacher à un autre groupe, soit dans l'Inde, soit hors de l'Inde, sont restées vaines.

L'aspect de la carte (v. la planche 9) et les vraisemblances historiques permettent d'imaginer que le dravidien occupait jadis un territoire plus vaste. Le fait le plus frappant est la position du brahui, isolé en plein Beloutchistan. Le brahui s'apparente plus particulièrement aux langues septentrionales de la famille dravidienne ; mais a-t-il été transporté au Beloutchistan par des émigrants de l'Inde centrale, de la même façon que le kurukh qui se parle aujourd'hui au Chota-Nagpour ? ou bien au contraire, s'autorisant des mœurs actuellement sédentaires des Brahui, et du fait que l'indo-aryen n'est pas indigène dans l'Inde, faut-il combler par la pensée l'énorme intervalle entre le Dekhan et le Beloutchistan, et admettre que jadis le dravidien se parlait dans toute l'Inde occidentale ? Questions impossibles à résoudre actuellement. Ce qui est sûr, c'est que le dravidien recule devant l'indo-aryen sous nos yeux mêmes. En effet, là où de grandes langues s'affrontent, comme le marathe (indo-aryen) et le canara (dravidien), on constate simplement sur la frontière commune des

influences réciproques dans le détail. Mais les langues moins civilisées sont aussi moins résistantes. Ainsi le groupe kurukh-malto, qui par ailleurs semble (ainsi que le télougou) empiéter sur le muṇḍā, a emprunté des mots et jusqu'à des éléments grammaticaux à l'indo-aryen. Le cas le plus frappant est celui du gondi, dont le territoire n'est qu'un ensemble d'îlots : ce parler doit lutter non seulement contre le télougou qui lui est apparenté, mais en outre contre diverses langues aryennes ; et les statistiques attestent que de 1891 à 1901 le nombre des sujets parlant cette langue était descendu de 1.300.000 à 1.100.000.

*
* *

Les langues dravidiennes sont parlées par près de 63 millions d'individus, c'est-à-dire environ le cinquième de la population totale de l'Inde. Parmi ces langues, le groupe de l'Inde centrale et le brahui sont des parlers dispersés, appartenant à des populations relativement peu civilisées ; les langues de la péninsule au contraire, télougou, canara, tamoul et malayalam, occupent une aire cohérente et ont un passé de culture qui remonte assez haut.

Le *tamoul*, qui est la plus connue d'entre elles, est parlé par plus de 18 millions d'hommes dans la plaine qui s'étend entre les Ghat orientaux et la côte du Coromandel, de Madras au cap Comorin (les territoires de Pondichéry et Karikal sont inclus dans ce domaine), et dans la plaine septentrionale de l'île de Ceylan ; les émigrants le transportent aussi dans les usines de Birmanie, les bazars d'Indochine et jusque dans les plantations des îles Fidji ou de l'Afrique australe. On doit annexer au tamoul le *malayalam*, qui est parlé par 6 à 7 millions d'hommes sur la côte du Malabar, du cap Comorin jusqu'à Cassergode (Kasargodu) [Mahé est sur cette côte]. Le malayalam est un dialecte du tamoul qui s'en est détaché tard, et qu'on distingue du tamoul parce qu'il a une littérature propre, depuis le XIIIe siècle au plus tôt.

La littérature tamoule est la plus ancienne et la plus riche des littératures dravidiennes, et même de toutes les littératures de

l'Inde, sauf la sanskrite. Les origines en sont légendaires, et l'ancienneté des œuvres qu'on possède, fort discutée. L'un des plus anciens textes est une grammaire, le *Tolkāppiyam* ou " ancien poème", où il est fait allusion à un alphabet probablement postérieur au IVe siècle de notre ère ; mais ce livre suppose une littérature antérieure, dont il fournit quelques extraits. Née sous la dépendance de la littérature sanskrite, la littérature tamoule contient cependant des œuvres d'une beauté originale, et en particulier une poésie gnomique raffinée ; le *Kuṛal* ,recueil de distiques attribué au paria Tiruvaḷḷuvar, en est un représentant célèbre (il en existe des traductions françaises). La langue littéraire classique est fort archaïque par rapport à la langue couramment parlée à présent, et un Tamoul sans éducation serait hors d'état de la comprendre. — La littérature en malayalam est pour la plus grande part imitée du tamoul ou du sanskrit.

Le tamoul est la langue dravidienne le plus anciennement connue en Europe. On avait d'abord donné à la famille entière le nom de " tamilien, tamoulien"; le nom de " dravidien ", choisi par Caldwell comme plus général, n'est en fin de compte qu'une forme plus ancienne du même nom, connue par le sanskrit.

Le *canara*, parlé par 10 millions 1/2 d'individus, est la langue du Maisour et de la partie sud-ouest des États du Nizam jusqu'à Bidar (environ 120 km. au Nord-Ouest de Haiderabad) ; de Bidar à Karwar (sur la côte, au Sud de Goa) il a une limite commune avec le marathe et le concani, dialectes indo-aryens, tandis qu'à l'Est il a pour voisins le tamoul et le télougou, tous deux dravidiens; enfin, le domaine canara occupe la côte entre Karvar et Mangalore. Le tuḷu, que 560.000 hommes parlent dans la région de Mangalore, est un dialecte très voisin du canara ; il en est de même pour les dialectes du plateau des Nilgiri en usage dans la société composite formée des agriculteurs *Baḍaga* (30.000), des artisans *Kota* (1280) et des *Toda* purement pasteurs, qui ne sont plus que 700 environ ; ces derniers sont particulièrement célèbres auprès des ethnographes, qui les ont étudiés avec d'autant plus de soin que cette tribu paraît vouée à une disparition prochaine.

Quoique le canara se parle sur une côte maritime, il reste confiné, sinon absolument dans son domaine propre, au moins dans l'Inde, où il compte quelques colonies isolées, dont une seule importante (plus de 100.000 à Madoura).

Le canara est la langue dravidienne le plus anciennement attestée de façon sûre : il se rencontre sur une courte inscription de la fin du ve siècle après J.-C. Comme pour le tamoul, les débuts de la littérature semblent dus aux Jains, et les premières œuvres sont largement influencées par le sanskrit et de style très raffiné. La plus anciennement datée est un art poétique du IXe siècle, qui mentionne des écrivains antérieurs (dont tous les noms sont sanskrits). On distingue d'après l'état de la langue trois périodes dans la littérature canara ; ces périodes correspondent aussi à la floraison des diverses sectes religieuses qui ont tour à tour inspiré la littérature.

Le *télougou* occupe la côte orientale, de Madras jusqu'au mont Mahendragiri, au Sud de Ganjam (Yanaon est sur cette côte) ; là son domaine confine à celui d'un dialecte indo-aryen, l'oriya, maître du delta de la Mahanadi. Plus à l'Ouest il rencontre encore un dialecte indo-aryen, le marathe. Entre les deux, il a comme voisin un dialecte dravidien, le gondi, dont certaines enclaves se trouvent sur son propre territoire. Au Sud et à l'Ouest aussi les langues voisines sont dravidiennes ; ce sont les grandes langues dont nous avons parlé, le tamoul et le canara ; le télougou est la langue dravidienne employée par le plus grand nombre d'individus : ce nombre atteint presque 24 millions ; il est parlé hors de son domaine propre et même hors de l'Inde, mais à un degré moindre que le tamoul.

Le premier écrivain télougou connu vivait au XIe siècle ; on connaît de lui une grammaire et une traduction du Mahābhārata ; ainsi la littérature en télougou débute comme les autres littératures dravidiennes sous l'influence sanskrite. Il faut signaler cependant qu'un savant indigène a noté chez cet écrivain et dans une inscription archaïque récemment découverte un mètre poétique qui serait emprunté à la littérature canara ; d'autre part les alphabets canara et télougou sont identiques depuis l'origine ; ainsi la

littérature télougou serait peut-être née sous une double inspiration.

Le groupe dravidien du Nord est constitué par une série d'îlots occupés par des tribus sans culture. La plus célèbre et la plus nombreuse est celle des Gond, qui ont donné leur nom au Gondwana qu'ils habitent. Les Gond sont une nation déchue ; leur langue a cédé et cède de plus en plus devant les langues plus civilisées, le marathe, l'hindi, l'oriya, et même sur certains points le télougou (ce dont la carte à elle seule suffirait à témoigner) ; sur trois millions de Gond que compte le recensement, la moitié au plus parle le *gondi* ; ce sont eux qui occupent les plateaux (sur l'un d'entre eux, les monts Mahadeo, ils voisinent avec la tribu muṇḍa la plus occidentale, celle des Kurku) ; ceux des plaines et des vallées ont perdu leurs langues en se civilisant.

Un autre groupe de parlers des mêmes régions, celui des *Kolam* et des *Bhil* du district de Basim (24.000 en tout), est très proche du dialecte gondi, mais rappelle par certains traits le canara et même le toda, ce qui fait supposer qu'il a eu une histoire particulière ; le kolami et le bhili sont eux aussi menacés de disparaître.

Le *kui*, parlé par les Ku (ou Khaṇd, ou Khond) établis sur les plateaux au Sud et au Nord de la trouée de la Mahanadi, est plus important que les précédents par le nombre (1 million 1/2), mais également inculte et menacé par l'oriya qui l'encercle. Plus au Nord encore, c'est toujours sur les montagnes qu'on rencontre des populations parlant un dialecte dravidien, et voisines de tribus muṇḍā : comme le gondi confine au kurku sur les monts Mahadeo, comme le kui voisine avec le savara dans les Ghat orientaux, ici le *kurukh* ou *oraon* du Chota-Nagpour et le *malto* du Rajmahal (65.000) se rencontrent avec le groupe principal du muṇḍā : mundari, kharia, korva, santali. Du reste ils sont ici de nouveaux arrivés : Kurukh et Maler sont originaires du Carnatic, et certains des villages qu'ils occupent ont des noms muṇḍā. Par leur uniformité les deux parlers témoignent de l'origine unique et du déplacement récent des tribus ; cependant le malto a subi, surtout dans son vocabulaire, de fortes influences de l'indo-aryen et même du santali.

Reste le *brahui*, parlé par 174.000 individus dans les montagnes du Beloutchistan oriental ; on a signalé plus haut l'intérêt de ce dialecte, qui du reste est lui aussi menacé dans son existence par les langues indo-européennes (indiennes et iraniennes) qui se parlent autour de lui.

*
* *

L'histoire du groupe dravidien est difficile à faire. La littérature télougou ne remonte pas au delà de l'an mille ; le plus ancien texte canara date des environs de l'an 500 ; la littérature tamoule remonte sans doute plus haut : mais tous les alphabets dravidiens dérivent d'alphabets de l'Inde septentrionale du IVe ou du Ve siècle. On sait par des témoignages antérieurs que l'Inde du Sud était peuplée et civilisée bien avant l'ère chrétienne : sans remonter aux stations préhistoriques dont les anthropologues ont seuls à tenir compte, on connaît les noms de plusieurs dynasties qui régnaient sûrement au IIIe siècle avant notre ère, et probablement avant, dans les différentes parties de l'Inde du Sud ; on a la preuve d'un commerce actif de cette région avec l'Europe vers l'époque du Christ, avec l'Extrême-Orient plus tôt encore ; mais sauf par quelques noms propres, on n'a aucun témoignage sur les langues qui y étaient parlées. Les noms de produits indigènes connus en Europe avant l'époque chrétienne sont ou peuvent être venus par des intermédiaires non-dravidiens. On a aussi cru trouver dans la Bible un nom du " paon " d'origine dravidienne ; mais le sens du mot hébreu en question est mal établi, et le mot tamoul auquel on le compare n'a pris le sens de « paon » que secondairement. Enfin on a essayé de reconnaître du canara dans quelques mots donnés comme indiens dans une farce grecque conservée sur un papyrus du IIe siècle ; mais les rapprochements proposés sont trop hasardeux pour qu'on puisse en faire état.

Ainsi là même où les langues dravidiennes ont un passé, l'histoire n'en remonte pas très haut. Et dans ces cas, l'état ancien diffère peu de l'état présent, surtout si l'on considère la grammaire. La conséquence est qu'il est impossible de se faire

une idée précise de la langue commune à laquelle remontent les langues dravidiennes actuellement parlées. Exception faite de quelques archaïsmes conservés dans la littérature, on ne peut guère définir le dravidien que par les caractéristiques communes aux langues actuelles du groupe. Ces caractéristiques sont assez nettes et assez constantes pour que la parenté de ces langues ait été aisément reconnue d'après leur aspect présent (pour le brahui la démonstration a été naturellement plus difficile) ; mais ces similitudes de structure peuvent être relativement récentes. L'aspect phonétique des langues dravidiennes n'est certainement pas ancien dans tous ses détails. Il est donc possible qu'une partie des caractéristiques grammaticales soit le résultat d'évolutions parallèles.

*
* *

Caractéristiques grammaticales.

LES SONS.

Le système phonétique des langues dravidiennes ressemble assez et a dû jadis ressembler davantage à celui de l'indo-aryen. Il s'en distingue d'abord par le fait que *e* et *o*, toujours longs en sanskrit, sont en dravidien brefs ou longs comme les autres voyelles. Parmi les consonnes, le dravidien possède une série inconnue aux autres langues de l'Inde ; cette série comprend en tamoul l'occlusive T [1] (forme sonore R), la nasale N, la liquide L ; R et L sont attestés en vieux canara, R seulement en vieux télougou ; les autres langues n'ont rien conservé. La prononciation ancienne de ces sons est mal connue ; il s'agit sans doute d'une série mouillée ; actuellement R est un « *r* fort », pratiquement confondu avec *r* ordinaire ; en tamoul sa forme géminée qui est sourde se prononce comme *tt* ou *ttr* mouillés à leur commencement ; L est une liquide cérébrale mouillée, qui n'a plus d'existence indépendante qu'en tamoul ; elle s'y prononce ẓ et

1. Dans ce chapitre, les petites capitales ne servent pas à transcrire des douces sourdes, mais exclusivement la série ici définie.

même *y* au Nord, tandis qu'au Sud elle s'est confondue avec *ḷ* cérébral (c'est ce qui est advenu aussi en canara) : de là vient que nous appelons « tamoul » le *tamiL*.

LA FLEXION.

Les racines dravidiennes peuvent être d'une forme quelconque. A l'état nu elles sont un nom au cas-sujet ou un verbe à l'impératif du singulier : tam. *aL* « pleure », *sol* « mot » ; *aḍi* « frappe » ou « pied », *uṇ* « mange » ou « nourriture ».

Il ne s'ajoute jamais d'éléments nouveaux en avant ou à l'intérieur de la racine. Le dravidien ignore la préfixation et l'infixation, et ne connaît que la suffixation ; le rôle des mots dans la phrase est donc marqué par les désinences.

NOM. — La flexion nominale est à la base de la morphologie entière ; c'est donc par elle qu'il convient de commencer. La flexion nominale à son tour est dominée par la catégorie du genre. Le substantif lui-même peut, il est vrai, ne comporter aucune caractéristique de genre ; mais le pronom et le verbe, qui s'accordent avec le substantif, en portent la marque dans leurs désinences.

Les grammairiens tamouls divisent les noms en noms « de haute classe » et « sans classe » ; en télougou on dit, en employant une expression sanskrite, « supérieurs » et « inférieurs », ce qui revient au même ; seules les grammaires du canara ont emprunté au sanskrit la classification courante en masculin, féminin et neutre. Les noms « supérieurs » désignent les dieux, les démons et les hommes ; les « inférieurs » désignent les animaux et les êtres inanimés.

Le traitement des femelles de « haute classe » varie. Il faut d'abord mettre à part, ici comme sur beaucoup d'autres points, le brahui, qui ne connaît plus le genre. Les autres langues se répartissent en plusieurs groupes. Le gondi et le kui rangent complètement déesses et femmes parmi les êtres « inférieurs » ;

le kurukh-malto, le kolami et le télougou font de même au singulier et les assimilent aux mâles correspondants, au pluriel seulement ; seules les grandes langues du Sud, tamoul-malayalam et canara, assimilent mâles et femelles « de haute classe » au pluriel et distinguent, mais au singulier seulement, les femelles, à la fois des mâles correspondants et des êtres « inférieurs ».

Au reste la distinction des sexes apparaît souvent dans le nom même (noms dérivés, pronoms, verbes) par l'affixation de particules qui ne sont autres que des mots primitivement indépendants, voulant dire « mâle », « fils », « femelle », « femme », « tante » (ou « belle-mère »). Mais ceci n'est pas proprement un procédé flexionnel.

On voit s'enchevêtrer ici plusieurs systèmes mentaux. Celui du gondi paraît plus primitif, et d'autre part l'état du tamoul et du canara correspond à leur degré de culture. On signale en kurukh un fait curieux : le pluriel des noms désignant des êtres femelles est du type « supérieur » dans une conversation entre hommes ou entre hommes et femmes ; si la conversation se tient entre femmes seulement, les noms en question sont traités comme au singulier, selon le type « inférieur » ; ainsi les femmes appliquent la règle du gondi.

Cette classification est inconnue dans l'Inde en dehors du dravidien. L'indo-aryen distingue ou a distingué trois genres, et le muṇḍā classe les êtres en animés et inanimés, ce qui est tout différent du système dravidien. — De même le muṇḍā a trois nombres, comme jadis l'indo-aryen : singulier, pluriel et duel. Le dravidien ne connaît que les deux premiers. La différence relative au nombre est du reste moins fondamentale que celle concernant le genre.

Comme on vient de le voir, la forme du pluriel des noms dépend de leur classe. Sauf en gondi-kui, il existe une désinence commune aux noms d'êtres supérieurs des deux sexes (généralement en *-r* : tamoul *maga*N « fils », pl. *magar* ; *tāy* « mère », pl. *tāyar* ; *nī* « toi », *nīr* « vous » ; kurukh *āl* « homme », pl. *ālar* ; *ās es'as* « il a brisé », *ār es'ar* « ils ont brisé »). Les noms

« inférieurs » n'ont souvent pas et n'avaient sans doute pas à l'origine de forme spéciale pour le pluriel ; aujourd'hui encore en tamoul courant on dira *nālu māḍu meygiʀadu* « quatre bœuf paît ». Il existe cependant une désinence assez répandue de neutre pluriel (en *-a*), d'origine incertaine du reste. Enfin on trouve un peu partout une désinence s'appliquant aux noms de tous genres, et s'ajoutant le cas échéant aux désinences déjà mentionnées : gondi et brahui *-k*, kui *-gā* ; can. *-gaḷu*, tél.*-lu* etc. ; celle-ci est sans doute d'extension récente.

Les formes jusqu'ici examinées ne sont que des formes de cas-sujet. Aux autres cas, les désinences sont les mêmes pour le singulier et pour le pluriel (tamoul sg. *maɴidaɴ* « homme », *maɴidargaḷ* « hommes » ; acc. *maɴidaɴei*, *maɴidargaḷei* ; gén. *maɴidaɴuḍeya*, *maɴidargaḷuḍeya* ; *vīḍu* « maison », *vīḍugaḷ* « maisons » ; *vīṭṭil* « dans la maison », *vīḍugaḷil* « dans les maisons » ; — kurukh *āl* « homme », plur. *ālar* : *ālgē* « à l'homme », *ālargē* « aux hommes »). Ce fait montre immédiatement que les désinences sont en réalité des mots indépendants postposés. Il en est en effet dont le sens est clair, surtout dans les cas à valeur concrète ; par exemple en tamoul *il* « maison » vaut « dans » lorsqu'il s'applique à un nom comme *ūr* « ville » : *ūril* « dans la ville » ; par suite il donne le sens conditionnel à une racine de sens verbal : *seyyil* « dans le faire, si l'on fait » ; de même on retrouve tél. *tōḍu* « aide », can. *oḍa* « union » dans tél. *-toḍan*, can. *-oḍan*, tam. *-ōḍu -uḍaɴ*, peut-être tulu *-ḍa* « avec » ; le complément du nom se marque en tamoul par *-uḍei* ou *-uḍeiya* qui veut dire « possession » ou « possédant ».

Ces postpositions s'ajoutent à une forme du nom qui est tantôt identique au cas-sujet, tantôt distincte et obtenue par suffixation. Cette seconde forme, dite « oblique », est en réalité une forme déclinée, susceptible d'avoir un sens par elle-même. Ainsi, en tamoul, de *kal* « pierre », la forme de régime direct est *kallei* ou *kalliɴei*, l'instrumental *kallōḍu* ou *kalliɴōḍu* ; or *kalliɴ* sert aussi bien de génitif que *kalliɴuḍeya* ; l'oblique prend ainsi valeur d'adjectif : *maleiyiɴ vaʟi* « chemin de montagne, montagneux ». Dans *taleiyiɴ* (*talei* « tête ») *iʟinda mayir* « cheveu tombé de la

tête » il donne un sens d'ablatif ; et c'est ainsi qu'il peut servir à indiquer la valeur comparative : *ada*NIT (pour -IN) *perid idu* « ceci [est plus] grand que cela ». Dans le pronom, l'oblique marque normalement la dépendance par rapport à un nom : tam. *nā*N « moi », *e*N*-a-kku* « à moi », *e*N *kaigal* « mes mains » ; *adu* « cela » *ada*N*-il* « en cela », *ada*N *vilei* « le prix de cela ».

PRONOM PERSONNEL. — Les pronoms personnels des deux premières personnes forment leur oblique par des alternances vocaliques. Comme le principe de ces alternances est par ailleurs presque inconnu au dravidien, on peut penser qu'il s'agit là de phénomènes secondaires dus au caractère accessoire de ces mots.

Ils n'ont pas de forme particulière à chaque sexe.

Le point le plus intéressant de la déclinaison des pronoms personnels est la formation du pluriel de la première personne. Le tamoul actuel, le télougou, le kui et le kurukh y distinguent deux formes, dites « inclusive » (« nous = moi et vous à qui je parle ») et « exclusive » (« nous autres, pas vous »). Le canara, le gondi, le brahui ne font pas cette distinction. On penserait qu'ils l'ont perdue, si certains textes, parmi les plus anciens, du tamoul ne l'ignoraient pas également. Certains auteurs considèrent dès lors la distinction des deux formes comme une innovation, et l'attribuent à une influence muṇḍā (le muṇḍā qui connaît le duel et le pluriel, distingue « nous deux = toi et moi », « nous autres deux », « toi et nous » et « nous autres (au moins trois) ». Mais une influence muṇḍā aussi généralisée, agissant sur des langues civilisées et lointaines, tandis que le gondi y aurait échappé, paraît peu vraisemblable. Il faut donc provisoirement se borner à enregistrer l'opposition entre le groupe oriental qui présente la distinction et le groupe occidental qui l'ignore.

VERBE. — Dans le verbe, les formes personnelles tendent partout à se rapprocher de la déclinaison pronominale. Même, à la 3[e] personne, elles apparaissent à première vue comme des noms dérivés au cas sujet. Rien, sauf le sens, ne distingue tam. *vill-a*N ou *villava*N « archer, » formés sur *vil* « arc », d'un nom d'agent

comme *sey-d-ava*N ou *sey-v-ā*N « faiseur », dérivés de *sey-* « acte » ; or *sey-v-ā*N vaut « il fera », *sey-d-ā*N « il a fait » ; de même en kurukh, *is'us* « le briseur », *ās es'as* « il a brisé ». Le pluriel se forme régulièrement : tam. *sey-v-ār* « (ils, elles) feront », malay. *seyvar*, comme *ivar* « ceux-ci » (pluriel de *iva*N), *manusar* « les hommes », *Piḷḷeyār* « l'Enfant » (pluriel honorifique) ; de même kurukh *ār es'ar* « ils ont brisé ». Et il en est du genre comme du nombre. La seule différence entre les deux formations tient à la présence ou à l'absence de suffixes temporels ; or il en est des noms d'action comme des noms d'agent : par exemple on retrouve le même suffixe du passé dans tam. *sey-d-al*, *sey-d-amei* « action d'avoir fait » en regard de *seyyal* « action de faire ».

D'autre part les désinences des deux premières personnes s'accommodent aux pronoms personnels : tam. *sey-d-ē*N « j'ai fait », en regard de vieux tam. *yā*N, oblique *e*N « moi » ; v. tam. *vā*L*u-nam* « nous vivons », en regard de *nām*, *nam* « nous » ; tam. *sey-d-īr* « vous avez fait », en regard de *nīr* « vous ». Ces désinences pouvaient en vieux tamoul s'ajouter à un nom quelconque : *ko*N*-ē*N (*kō*N « roi ») « je suis roi », *sāmi-(y)īr* « vous êtes maîtres ». On retrouve entre ces formes et les précédentes la même parenté qu'entre le nom d'agent et les formes verbales de 3e personne. Elles peuvent être accompagnées d'un adjectif ; par exemple *perum pūṇē*N « j'ai un grand ornement » (*pūṇ* « ornement ») et être elles-mêmes déclinées : *perum pūṇē*N*ukku* « à moi qui ai un grand ornement ».

Le sens verbal dans une telle formation ne résulte en somme que du sens propre du mot et de son rôle dans la phrase.

Mais le parallélisme entre la conjugaison et la déclinaison des noms et des pronoms ne se poursuit pas dans tous les détails. Certaines désinences restent isolées : p. ex. tam. *nī varu-gi*R*-āy* « tu viens », kurukh *nīn es-k-ai* « tu (masc.) as brisé » ; ou tam. *a*R*i-ba* « ils savent », *kaṇ b-al* « je verrai ». Ailleurs l'assimilation incomplète laisse soupçonner l'existence de désinences plus anciennes : tam. *nām seydēm* ou *seydōm* « nous avons fait » ; can. *nā(nu) māḍuttēne* « je fais », *nāvu māḍuttēve* « nous faisons ». Il faut enfin noter la présence de certaines formations indécli-

nables, par ex. tamoul *-um* : *seyyum* « cela fera », *avaḷ varum* « elle viendra », *aḷikkum vānaravīrar* « les héros singes accorderont », *añjudum* « nous avons craint » ; dans *suvāmi vandadu* « le seigneur est venu », le verbe est au neutre.

L'aspect nominal du verbe est donc probablement un fait secondaire, résultant d'une tendance commune à tout le dravidien ; la langue d'où dérivent toutes les langues actuelles séparait peut-être nettement le verbe du nom. Mais si la tendance à l'assimilation du verbe au nom est comparativement récente, elle ne s'explique pas par des influences non-dravidiennes ; l'indo-européen de l'Inde n'a rien de comparable ; en muṇḍā, le nom d'action n'est pas fléchi, et si l'on fait usage en muṇḍā de suffixes pronominaux dans le cas d'êtres animés, c'est en les accolant au sujet et non pas au nom d'action.

Il résulte de la formation des désinences sus-indiquée que le verbe dravidien ne comporte ni voix ni modes ; ces nuances, comme celles d'aspect et de temps, se marquent par les suffixes ou par des périphrases.

On dit généralement que le dravidien possède un verbe négatif. Ceci n'est pas universel ; ainsi le kurukh a des négations du type ordinaire et un verbe signifiant « n'être pas », ce qui est pure affaire de vocabulaire. Le procédé le plus général est la suffixation, selon le type normal. Par exemple le tamoul fait usage d'un suffixe *-al* ou *-il* (p. ex. *sey-d-il-ei* « tu n'as pas fait ») qui par ailleurs est susceptible de servir à la dérivation nominale (*irei-y-il-i* « exempt d'impôt ») ou de prendre lui-même la forme verbale (*ivaɴ kolei-ppaḍu magaɴ alaɴ* « celui-ci n'est pas un homme à tuer »). Le procédé le plus curieux est l'emploi d'un suffixe *zéro* en tamoul, canara et gondi ; le négatif s'obtient par la suppression de tout signe d'aspect ou de temps : tam. *pēs-ēɴ* « je ne parlerai pas », can. *māḍe* « je ne fais pas », gondi *tinnōn* « je ne mange pas » (en regard de *tindātōnā* « je mange », *tindākā* « je mangerai », *tittōnā* « j'ai mangé » etc.) ; selon Caldwell cette forme étrange ne serait qu'un aspect, dû à une altération phonétique, du suffixe *-a-* attesté avec le sens négatif en télougou, en kui et en brahui ; selon Kittel au contraire, il s'agirait d'une forme à sens primitivement éventuel.

Phrase.

La phrase dravidienne apparaît donc dans l'ensemble comme une phrase nominale où le prédicat serait une sorte de participe capable de rection ; ce prédicat est toujours à la fin de la phrase. Ex. tamoul : *seṭṭi-y-iN maga*N « le fils du marchand » *nē*TT*u* « hier » *periya vīṭṭei* « la grande [indéclinable ; il n'y a pas d'adjectif au sens indo-européen] maison [acc.] » *kaṇḍā*N « a vu ».

Il n'y a pas de subordination (ni de pronom relatif) ; la place en est prise :

1° par des formes nominales, fléchies ; ex. : *nā*N *vari*N « si je viens » : -*i*N est une désinence d'oblique nominal ; mais on ne peut traduire par « à la venue de moi », car le pronom *nā*N est au cas sujet, comme s'il précédait un verbe.

2° par des formes nominales non fléchies ; exemples :

*avaḷ vanda po*L*udu* ;
« elle venant (au) moment », « quand elle vient » ;
*nā*N *seyd uḷi* ;
« moi (sujet) fait (au) lieu », « pendant que je faisais ».

3° par des gérondifs indéclinables variés ; exemples :

*kāṇa vammi*N « venez voir » ;
*appo*L*udu · ava*N *kopam aḍeindu*
« (à) ce moment lui colère ayant pris
*uḷḷe pōga ma*N*ad illād irundā*N ;
dedans aller (l')idée n'étant pas était »,
« alors s'étant mis en colère, il ne voulait pas entrer » ;
*maram uṇḍ e*ND*u kaṇgi*RēN
« (un) arbre existe ayant dit je vois »,
« je vois qu'un arbre existe » (cette manière de rendre le « que » du français par une forme signifiant « ayant dit » est fréquente dans l'Inde, même dans les langues aryennes).

J. Bloch.

INDICATIONS BIBLIOGRAPHIQUES

La grammaire comparée du dravidien a été fondée en 1856 par Caldwell, évêque de Tinnevelly ; son livre, *A comparative grammar of the dravidian or south-indian family of languages*, paru d'abord en 1856, a été refondu en 1875 ; la troisième édition, publiée longtemps après sa mort (en 1913) est une réimpression, abrégée et nullement modifiée, de la seconde. Le livre de Caldwell devrait être refait, mais c'est en somme le seul qu'on ait sur la matière, et tout ce qui a été écrit depuis en dérive.

On trouvera un bon résumé d'ensemble, une bibliographie complète et un grand nombre de faits nouveaux dans la partie du *Linguistic Survey of India* de M. Grierson consacrée au dravidien, qui est l'œuvre de M. Sten Konow (vol. IV, p. 277-681 ; Calcutta, 1906).

En 1919 a commencé de paraître sous les auspices de l'Université de Madras et sous la direction de M. Collins, une série de *Dravidic studies*, dont trois fascicules ont paru jusqu'à présent ; l'un des auteurs, M. K. V. Subbaya, a aussi publié dans l'*Indian Antiquary*, de 1909 à 1911, quelques articles comparatifs utiles.

Parmi les grammaires de langues particulières parues depuis la publication du *Survey*, il faut signaler surtout *The brahui language* de M. Denys Bray (Calcutta, 1909).

Le livre de Rivers consacré à l'étude ethnographique des Toda (1906) contient sur leur parler des documents précieux.

En français, le *Manuel de langue tamoule* de M. Vinson (Paris, 1903) est riche en faits de détail.

LE SINO-TIBÉTAIN

GÉNÉRALITÉS

On groupe généralement en une seule famille les langues chinoise, tai et tibéto-birmanes (voir planche 10). Nous adoptons ici cette classification qui permet de faire ressortir les affinités de ces grandes masses linguistiques. Toutefois il convient d'observer dès l'abord qu'il ne peut être question de faire actuellement la théorie du sino-tibétain commun. La grammaire comparée des langues tai et des langues tibéto-birmanes n'est pas encore faite. D'autre part la comparaison des dialectes chinois modernes et les indications fournies par les anciens lexiques ne permettent de reconstituer le chinois commun que pour une époque où cette langue était déjà fort usée et très loin de ses origines. Dans ces conditions, il est impossible de savoir, avec certitude et précision, ce qu'était l'ancêtre dont semblent issus le chinois archaïque, le tai commun et le tibéto-birman commun. Tout ce qu'on peut faire est d'indiquer les rapports de ces langues diverses et d'énumérer par là même les raisons qu'on a de les rapprocher.

En chinois, pas plus qu'en tai ou en tibéto-birman, nous ne pouvons nous flatter d'isoler sûrement les racines primitives. Il est probable qu'à date ancienne, dans toutes ces langues, de nombreux mots étaient plus longs qu'ils ne le sont aujourd'hui et comprenaient outre la racine, un ou plusieurs affixes et peut-être même une désinence. Au cours des siècles, ces agrégats se sont réduits graduellement. L'usure a principalement atteint la fin et le début des mots, mais nous n'avons aucune raison d'admettre

que seuls aient disparu les affixes et désinences laissant intacte une racine présumée monosyllabique. Les parties les moins résistantes ont dû s'effriter, qu'elles appartinssent ou non à la racine ; il en est résulté des contractions et des altérations diverses, et les mots ainsi raccourcis ne sont, dans certains cas, que des débris très différents des racines originelles.

On voit combien il serait inexact de définir le sino-tibétain commun « une langue monosyllabique » si l'on entend par là un idiome dont tous les mots dès l'origine n'auraient eu qu'une seule syllabe. Tout au plus peut-on parler d'une tendance au monosyllabisme dans les langues issues de l'ancêtre commun. Cette tendance s'est d'ailleurs inégalement manifestée dans les diverses langues de la famille. Certains parlers tibéto-birmans ont conservé jusqu'à ce jour des préfixes syllabiques du type *ba*, *da*, *ga*, *sa*. Mais dès le VIIe siècle, le tibétain ne possédait plus que des préfixes asyllabiques du type *b*, *d*, *g*, *s*. Parmi les langues tai, l'āhom a conservé jusqu'à l'époque moderne une assez grande variété de préfixes asyllabiques. Par contre, les préfixes avaient presque complètement disparu en chinois dès avant le VIe siècle.

D'une manière générale, on peut dire que l'usure des langues a été d'autant plus rapide que les peuples qui les parlaient jouaient dans l'histoire un rôle plus actif. Elle a été relativement lente chez certaines tribus himalayennes et assamaises qui vivaient à l'écart dans les hautes vallées sans se mêler aux hordes voisines. Elle a été plus rapide chez les Tibétains qui, dès les premiers siècles de notre ère, entrèrent étroitement en contact avec les autres peuples de l'Asie Centrale. Elle s'est exercée avec le maximum d'intensité chez les Chinois qui, pendant leur longue histoire, n'ont cessé de subir l'action des peuples étrangers et de réagir à leur tour.

Le système des affixes étant aujourd'hui ruiné dans la plupart des langues sino-tibétaines, les sujets parlants n'ont plus conscience des anciens procédés de dérivation et par suite du lien de parenté qui unit entre eux les divers mots issus d'une commune racine. Il en résulte en outre que les mots ont perdu presque entièrement leur souplesse et leur fluidité. Tandis qu'à l'époque ancienne,

des racines étaient susceptibles de prendre des sens variés en s'agrégeant des morphèmes et sans doute aussi en changeant de forme, il n'y a plus actuellement dans les langues les plus évoluées que des mots inertes. Cependant, comme il est toujours nécessaire d'indiquer les rapports des mots dans la phrase, on utilise à cet effet des « particules » indépendantes qui jouent le rôle des anciens morphèmes et qui sont, autant qu'on en peut juger, d'anciens mots réduits à marquer des relations grammaticales. Ces mots, employés à tout propos, se sont naturellement usés plus vite que les autres ; ils ont presque entièrement perdu leur ancien sens et leur puissance évocatrice et c'est pourquoi ils méritent le nom de « mots vides », que les grammairiens chinois leur ont donné.

Il n'est pas douteux qu'en sino-tibétain commun les racines pouvaient s'adjoindre des morphèmes ; elles devaient même être susceptibles de changer de forme par le jeu de certaines alternances. C'est du moins ce que paraît attester la conjugaison du verbe tibétain. Tandis qu'en chinois et en siamois, le verbe conserve toujours la même forme quels que soient le temps et le mode, ces circonstances étant marquées par des auxiliaires ou des particules, un bon nombre de verbes tibétains présentent des alternances variées au présent, au parfait, au futur, à l'impératif. Ex. : verbe « remplir »

Présent *'geṅs* Parf. *b-kaṅ* Fut. *d-gaṅ* Imp. *k'oṅ*

Certains auteurs, et notamment M. Conrady, prétendent qu'il n'y a point là une véritable conjugaison, parce que, dans les plus vieux textes tibétains, la spécialisation des formes verbales est imparfaite, telle modalité pouvant servir à rendre des temps différents. Il semble au contraire que les irrégularités mêmes du système sont une preuve de son ancienneté et que la conjugaison serait plus simple et plus régulière si elle était due, comme le pense M. Conrady, à la spécialisation arbitraire de formes primitivement équivalentes.

Si les variations du verbe tibétain ressemblent fort à une conjugaison, certaines alternances conservées ailleurs, dans les formes pronominales, font songer à une déclinaison. M. Karlgren a mon-

tré qu'en chinois archaïque les pronoms personnels de la 1re et 2e personne avaient des formes différentes suivant qu'ils étaient employés au cas-sujet ou au cas-régime.

	1re personne	2e personne
Nominatif génitif	**nguo*	**ńźi^wo*
Régime	**nga*	**ńźia*

En birman, les mêmes pronoms varient également suivant qu'ils sont au cas-sujet ou au cas-régime.

	1re personne	2e personne
Sujet	*nga-ga°*	*nin-ga°*
Régime	*nga°-go*	*nin°-go*

Le petit cercle placé à côté du mot sert à noter dans la transcription birmane le « check-tone », c'est-à-dire un ton analogue au « ton rentrant » du chinois. L'analogie entre les faits chinois et birmans est d'autant plus remarquable que le pronom chinois de la 1re personne correspond exactement au pronom tibéto-birman : tibétain *nga*, birman *nga*. Le pronom de la 2e personne singulier avait en chinois ancien plusieurs prononciations dont une au moins **ni* répond aux formes birmanes *nin* et *nyi*. Aujourd'hui encore en chinois moderne, le pronom de la 2e personne est *nin* ou *ni-na* au pluriel et au singulier respectueux.

Phonétiquement, la tendance à assourdir les anciennes initiales sonores est un trait commun aux langues chinoises, tai et tibéto-birmanes. M. Conrady a cherché à expliquer ce phénomène par la présence d'anciens préfixes. Il admet que le changement des anciennes sonores en sourdes est dû, dans tous les cas, à la fusion d'un préfixe avec la racine. Cette hypothèse n'a pas été confirmée par les recherches plus récentes. On sait aujourd'hui que les préfixes chinois avaient presque complètement disparu dès avant le VIe siècle et que l'assourdissement des initiales sonores est un phénomène beaucoup plus tardif. On ne peut concevoir que ces initiales aient été modifiées par les préfixes plusieurs siècles après la disparition de ces derniers.

Un autre caractère éminent des langues chinoise, tai et tibéto-

birmanes est qu'elles possédaient autrefois un système de tons qui a subsisté, au moins en partie, dans les langues modernes. Le principe fondamental du système ancien est que la hauteur du ton était en relation avec la nature de l'initiale : un mot commençant par une sourde était prononcé sur un ton plus élevé qu'un mot commençant par une sonore. Cette loi qui se vérifie en chinois ancien, en tai et en tibéto-birman constitue l'un des principaux arguments en faveur de la parenté de ces langues.

D'autre part, en chinois, en tai et en tibéto-birman, une même racine avait fréquemment une forme causative ou transitive qui s'opposait à une forme intransitive. Ces deux aspects différaient l'un de l'autre soit par la présence ou l'absence d'un préfixe, soit par la nature de l'initiale, sourde dans le premier cas et sonore dans le second. Cette question a été étudiée par M. Conrady dans un ouvrage qui, bien que caduc en plusieurs de ses parties, n'en demeure pas moins la plus importante contribution qu'on ait apportée jusqu'ici à l'étude des rapports du chinois, du tai et du tibéto-birman (*Eine indochinesische Causativ-Denominativ-Bildung und ihr Zusammenhang mit den Tonaccenten*, Leipzig, 1896).

Enfin ces trois groupes linguistiques ont en commun de nombreux éléments de vocabulaire, noms et verbes usuels, mots grammaticaux, noms de nombre.

Malgré les traits de ressemblance qui viennent d'être énumérés, il s'en faut que la généalogie de ces langues soit fixée d'une manière indiscutable et définitive. Sans doute, on a réussi à reconstituer le chinois ancien et on peut d'ores et déjà considérer comme établie l'existence d'une famille tai et d'une famille tibéto-birmane. Mais, sur chacun de ces domaines, les études de morphologie sont encore très peu avancées, en sorte que toute classification d'ensemble est provisoire. Si la parenté du chinois et du tibéto-birman nous paraît probable, ce n'est pas sans hésitation que nous y rattachons les langues tai. On verra plus loin que ces dernières, bien qu'encombrées de mots chinois, ont certaines affinités avec la famille mon-khmer. Il se peut que le progrès des connaissances nous oblige quelque jour à classer le tai commun dans la famille des langues austroasiatiques.

LES LANGUES TIBÉTO-BIRMANES

Les langues tibéto-birmanes sont parlées par environ 20 millions d'individus. Ceux-ci peuplent le Tibet et une grande partie de la Birmanie ; ils débordent en outre sur le versant sud de la chaîne himalayenne et forment quelques îlots dans la Chine du Sud et le Haut-Tonkin. Leur habitat primitif était sans doute beaucoup plus septentrional : les Tibétains qui, dès avant notre ère, étaient établis dans les montagnes du Kan-sou méridional, se sont progressivement déplacés vers le Sud, tandis que les Birmans suivaient probablement le cours de la Salouen et de l'Iraouaddy. En Birmanie et sur le versant sud de l'Himalaya, la langue des envahisseurs s'est substituée en partie aux parlers mon-khmer et muṇḍā précédemment en usage dans ces régions. Dans l'état actuel de nos connaissances, il est difficile de décider si tous les parlers considérés ici même comme tibéto-birmans appartiennent réellement à cette famille, ou si quelques-uns ne sont pas les restes d'anciens idiomes recouverts d'apports étrangers.

La famille tibéto-birmane comprend, comme le nom l'indique, deux importantes langues de civilisation : le tibétain et le birman. Le tibétain, dont l'ancienne littérature abonde en ouvrages bouddhiques traduits des langues indo-aryennes, a été écrit dès le VIIe siècle ap. J.-C. Le fameux traité tibéto-chinois gravé sur le pilier bilingue de Lhasa est de 822. D'autres langues parlées dans la région himalayenne, tels le nĕwārī et le leptcha, ont également une littérature assez importante.

L'écriture birmane est moins ancienne que l'écriture tibétaine. Une des plus vieilles inscriptions birmanes est gravée sur une des faces du pilier quadrilingue de Myazedi (1084 A. D. ?). Sur les quatre faces de ce pilier, le même texte est reproduit en pali,

en birman, en mon et en une quatrième langue naguère encore inconnue. M. Blagden, qui a réussi à déchiffrer cette dernière, admet qu'elle appartient à la famille tibéto-birmane et qu'elle était parlée par un ancien peuple de la région de Prome : les Pyu. Cette langue est connue au total par une quinzaine d'inscriptions dont quelques-unes seraient du VIIᵉ siècle de notre ère. Les inscriptions en birman et en pyu sont reproduites et étudiées dans *Epigraphia Birmanica*.

Le tibétain, le birman et les autres langues du Tibet et de Birmanie qui possèdent une littérature sont écrits au moyen d'alphabets empruntés à l'Inde. Par contre, au Kan-sou, c'est avec des caractères imités de l'écriture chinoise que s'écrivait la langue des Si-hia. Cette langue, oubliée depuis longtemps et qui nous est très imparfaitement connue, paraît appartenir à la famille tibéto-birmane. L'expression Si-hia ne désigne pas un peuple, mais une dynastie. Le système d'écriture qui porte ce nom fut constitué en 1037.

On peut diviser en quatre groupes la famille tibéto-birmane :

I. Groupe tibétain.
II. Groupe bodo-nāgā-katchin.
III. Groupe birman.
IV. Groupe lo-lo.

Dans chacun de ces groupes, on peut distinguer les sous-groupes, langues et dialectes suivants [1] :

I. — GROUPE TIBÉTAIN.

A. *Tibétain*, qui comprend : tibétain central, baltī, pūrik, ladakhī, lahul, spiti, nyamkat, jaḍ, garhwal, kāgate, sharpa, ḍānjongkā (Sikkim), lokhe (Bhutan), khams.

B. *Langues tibéto-himalayennes* : gurung, murmi, sunwār, magarī, nēwārī, pahrī, leptcha ou rong, ṭōṭō.

1. De nombreuses langues des trois premiers groupes n'étant guère connues que par des travaux anglais, on a conservé, sans modification, les noms qui désignent les parlers tibéto-birmans dans le *Linguistic Survey of India*.

C. *Langues du Nord-Assam* : aka ou hrusso, abor-miri, daflā, mishmi, digāru, mījū.

II. — GROUPE BODO-NĀGĀ-KATCHIN.

A. — *Bāṛā* ou *Bodo*. Ce sous-groupe comprend le bāṛā proprement dit, parlé par les mech et kachārī, le rābhā, lālung, dīmāsā ou kachārī des collines, gārō ou māndē, tipurā, chutiyā.

B. — *Nāgā*. Ce sous-groupe comprend au centre : āo ou hatigorria, lhōtā, tengsa-nāgā, thukumi et yachumi ; à l'Ouest : angāmi ou tengimā, simi ou semā, rengmā ou unzâ, kezhāmā ; à l'Est : angwānku ou tableng, tamlu ou chingmēgnu, banparā, chang ou mojung, assiringiā, mutoniā, mohongiā, namsangiā ; et enfin, un certain nombre de dialectes dont le classement est douteux : mikir, ēmpéo ou kachchā-nāgā, kabui ou kapwī, khoirāo, sopvomā ou māo-nāgā, marām, miyāngkhāng, kwoireng ou līyāng, luhūpā ou luppā, tāngkhul, phadāng, khangoi, maring.

C. — *Katchin*. La langue katchin ou singphō ou kakhyen est parlée par de nombreuses tribus disséminées depuis l'Assam jusqu'en Chine.

III. — GROUPE BIRMAN.

A. — *Kuki-tchin*. Ce sous-groupe comprend le meithei ou manipurī, et le kukitchin proprement dit, divisé en un très grand nombre de dialectes : thādo, soktē, siyin, rāltē, paitē, zahao ou yahow, shunkla ou tashōn, lai, lakher, lushēī, banjōgī, pānkhū, rāngkhōl, bētē, hallām, langrong, aimol, chiru, kolrēn ou koireng, kōm, kyau ou chāw, mhār, pūrūm, anāl, hirōi-lamgāng, khyang ou shö et khami.

B. — *Birman*. Le birman se divise également en un grand nombre de dialectes très mal connus parmi lesquels : l'arakanais ou maghī, le mrū, les dialectes khyaung-tha, yabaing, tavoy, myelat ou taungyo.

IV. — GROUPE LO-LO.

On peut ranger provisoirement dans ce groupe les langues tibéto-birmanes suivantes, parlées principalement sur les territoires de la Chine et du Haut-Tonkin : lo-lo, mo-so, si-hia, hu-ni, pu-la.

On trouvera des indications bibliographiques sur les trois premiers groupes dans *Linguistic Survey of India*, vol. III, sur le groupe lo-lo dans un article de B. Laufer, *The Si-hia language*, T'oung Pao 1916, et enfin sur les deux derniers groupes dans H. Cordier, *Bibliotheca Indosinica*.

Aux confins occidentaux du domaine tibétain, le burušaskī ou khajuna est parlé par quelques tribus des montagnes du Dardistan. Cette langue n'a pu être rattachée à aucun groupe linguistique.

On a déjà signalé plus haut la tendance à l'assourdissement des occlusives sonores qui se manifeste dans un certain nombre de langues tibéto-birmanes. Ainsi le tibétain classique a conservé les anciennes sonores, mais celles-ci sont devenues sourdes dans un certain nombre de dialectes modernes, notamment au Tibet central.

En tibétain classique, les préfixes sont généralement asyllabiques, c'est-à-dire qu'ils ont perdu leur élément vocalique. Mais, tandis qu'au Ladak ces préfixes asyllabiques sont encore prononcés distinctement, ils ont cessé de l'être dans les dialectes orientaux. Par contre, dans certaines langues de la même famille, les préfixes ont conservé leur voyelle et constituent une syllabe distincte.

Ex. « pousser »	tib. *g-noṅ-pa*	singphō	*ga-noṅ*.
« nez »	tib. *s-na*	lu-tse	*se-na*.
« tigre »	tib. *s-tag*	leptcha	*sa-t'aṅ*.

Certaines langues comme le katcharī ont même deux séries de préfixes à initiale aspirée et non aspirée.

ga da ba
k'a t'a p'a

Les anciens préfixes syllabiques se sont donc réduits dans les langues les plus évoluées de la famille tibéto-birmane. L'étude des suffixes permet d'entrevoir des changements analogues, bien qu'ici l'évolution semble avoir été moins rapide. Le tibétain possède encore un certain nombre de suffixes syllabiques parfaitement conservés. Toutefois, dans certains cas, il est visible qu'un mot simple en apparence résulte en fait de la fusion d'une racine et d'un suffixe. Par exemple, *rde'u* « petite pierre » provient apparemment de **rdo-bu*, c'est-à-dire d'une racine *rdo* « pierre » et du suffixe *bu* qui forme des diminutifs.

En somme, de nombreux mots tibétains réduits actuellement à une seule syllabe résultent de la fusion d'une racine et d'éléments syllabiques préfixés et suffixés. Ces mots sont souvent plus usés, plus courts que ceux des parlers « sauvages ». Il est clair qu'avant de prétendre approfondir la phonétique et la morphologie du tibéto-birman, il faudrait d'abord restituer aux mots leurs anciens éléments et leur forme polysyllabique. Ce travail n'a pas été fait ; on ignore la forme exacte des affixes primitifs et à plus forte raison leur fonction sémantique et grammaticale.

D'ores et déjà pourtant un point paraît acquis : certains préfixes, notamment celui qui a donné naissance à *s* tibétain, servaient à former des verbes transitifs et causatifs. On a, en tibétain :

nad-pa « malade » et *s-nad-pa* « blesser »,
riṅ-pa « long » et *s-riṅ-pa* « allonger », etc.

Parallèlement, on constate l'alternance suivante en tibétain et dans de nombreuses langues tibéto-birmanes : un mot à initiale sonore a un sens intransitif tandis que le même mot avec initiale sourde a un sens transitif ou causatif.

Par exemple en tibétain : *gaṅ* « plein » *b-kaṅ* parf. du verbe « remplir », *k'oṅ* impér. du même verbe.

Il est vrai que le même verbe a un futur *d-gaṅ* et un infinitif *'geṅs-pa* avec initiale sonore. Dans ce cas comme dans beaucoup d'exemples analogues, il n'est donc pas certain que l'alternance sonore/sourde soit directement en relation avec la formation d'un verbe causatif et sur ce point les opinions de M. Conrady ne doivent pas être adoptées sans réserve (voir ci-dessus, p. 364).

Du moins des faits comme l'alternance qui vient d'être signalée, et la série tibétaine :

nad « maladie » *nad-pa* « malade » *s-nad-pa* « blesser »
suffisent à prouver qu'en tibéto-birman une même racine a pu donner naissance à diverses catégories de mots : noms, adjectifs, verbes, ceux-ci étant différenciés par des affixes et par certaines modifications de la racine.

Dans les langues modernes tibéto-birmanes, les rapports des mots sont marqués par la place qu'ils occupent dans la phrase et par l'emploi de particules qui sont d'anciens mots spécialisés dans une fonction grammaticale. Les parties du discours se suivent normalement dans l'ordre suivant : sujet — régime — verbe. Le nom au génitif précède le nom qui le régit, soit immédiatement, soit par l'intermédiaire d'une particule.

Le substantif est en principe invariable. Les formes verbales sont beaucoup moins rigides. Les modes et les temps du verbe sont exprimés non seulement au moyen d'affixes, de particules et d'auxiliaires, mais aussi par d'importantes modifications de la racine. L'exemple déjà cité :

« remplir » *'geṅs-pa* Parf. *b-kaṅ* Fut. *d-gaṅ* Imp. *k'oṅ*
montre le rôle des affixes et le jeu des alternances : modification de la voyelle d'une part et de l'initiale d'autre part. M. Conrady a essayé de ramener toutes ces variations à une seule : la présence d'un préfixe aurait suffi à modifier l'initiale et par conséquent le ton puisque le ton était jadis en fonction de l'initiale. Cette hypothèse est probablement beaucoup trop simple : il est douteux que l'assourdissement des sonores initiales soit uniquement dû à l'influence des préfixes ; en tout cas, cette influence ne rend pas compte des alternances vocaliques.

La conjugaison a tendu de bonne heure à se simplifier. Déjà dans les plus vieux textes tibétains, les temps des verbes sont parfois confondus et les formes proprement verbales sont en régression, tandis qu'on emploie de préférence des formes nominales, sortes de participes et de gérondifs invariables. Au lieu de : « je frappe » on dit « par moi frappé ». On observe une évolution analogue dans un grand nombre de langues, notamment en indo-aryen et en iranien.

Si les substantifs n'ont pas la plasticité des formes verbales, certains indices donnent à penser que le jeu des alternances n'était pas limité jadis à la conjugaison des verbes. J'ai déjà signalé plus haut (p. 364), les variations des pronoms personnels birmans suivant qu'ils sont au cas-sujet ou au cas-régime. Dans le même ordre d'idées, certains dialectes du Nord-Assam présentent des faits aussi nets :

	PRONOM PERSONNEL 1e personne		PRONOM PERSONNEL 2e personne		PRONOM PERSONNEL 3e personne	
	Miri	Daflā	Miri	Daflā	Miri	Daflā
Cas-sujet	ngâ	ngâ	nâ	nâ	bui	ma
Cas-régime	ngōm	ngām	nōm	nām	buim	mām

Pour les substantifs désignant des êtres animés, le cas-régime est marqué en miri par la désinence *em*, en daflā par la désinence *am*. Pareillement en rong, la désinence *m* s'ajoute aux pronoms pour former un accusatif. Il est possible qu'en miri, l'alternance *ngâ/ngōm* s'explique par la fusion d'une désinence avec la racine. Le fait n'en vaut pas moins d'être noté.

En somme, à mesure qu'on remonte vers les origines, la langue apparaît plus riche et plus vivante ; les racines prennent des formes diverses suivant le sens à exprimer ; les mots s'adaptent dans une certaine mesure à leur fonction grammaticale. Au contraire, dans les langues modernes, la morphologie s'appauvrit, les racines perdent leur plasticité, les formes verbales s'éliminent, et il ne reste guère que des mots invariables dont les rapports s'expriment principalement au moyen de particules qui sont elles-mêmes d'anciens mots usés qu'un fréquent usage a vidés de leur contenu.

LE CHINOIS

Le chinois est une très importante langue de civilisation. Il est parlé sur un territoire plus vaste que l'Europe par une population qui ne doit pas compter moins de 300 millions d'individus. Il a été noté par l'écriture dès une antiquité très reculée. Malgré son imprécision et sa complication, cette écriture est encore en usage en Chine et dans les pays voisins. Par suite de l'expansion de la civilisation chinoise, la langue et l'écriture des civilisateurs se sont répandues notamment en Annam, en Corée et au Japon, et dans ces pays se sont conservées, au moins chez les hommes instruits, trois prononciations traditionnelles de la langue écrite dont la connaissance est précieuse pour l'étude de la vieille langue : le sino-annamite, le sino-coréen et le sino-japonais. En Chine, après des périodes de centralisation où tendait à se constituer une langue commune parlée dans tout l'empire, il se développait de nombreux dialectes. Actuellement, la langue mandarine où *kouan houa* est celle des fonctionnaires, et son usage s'est généralisé dans un certain nombre de provinces ; partout ailleurs, chaque district a son dialecte particulier et deux individus parlant des dialectes différents sont souvent incapables de se comprendre.

Chaque mot est représenté par un caractère distinct. Parmi les très nombreux caractères chinois, il n'y en a guère que 4.000 qui soient vraiment usuels. Pour les classer, on a choisi un certain nombre d'éléments graphiques appelés « clefs » ou « radicaux » qui sont communs à tout un groupe de caractères. Ces clefs, au nombre de 214, jouent dans les dictionnaires chinois le rôle des initiales dans les nôtres.

Certains caractères sont purement idéographiques. Les autres sont entièrement ou partiellement phonétiques : une « phonétique » chinoise est un signe qui, au lieu de figurer un objet ou une idée, a pour but de représenter la prononciation d'un mot.

On possède de très vieux textes écrits de cette manière. Sans doute les livres canoniques ne sont pas aussi anciens qu'on s'était plu à le supposer. Ils ont presque tous été remaniés à une époque voisine du début de l'ère chrétienne. Mais on a exhumé des os et des écailles de tortue inscrits qui datent du deuxième millénaire avant notre ère. Ces documents, de même que les livres les plus anciens, ne nous renseignent que d'une manière insuffisante sur la morphologie et la phonétique du chinois archaïque en raison de l'imprécision de l'écriture.

Dans l'état actuel de nos connaissances, c'est seulement pendant la période comprise entre le VI[e] siècle de notre ère et les temps modernes qu'on peut suivre l'évolution phonétique et morphologique du chinois. De récents travaux, fondés principalement sur les traités des anciens lexicographes et sur la comparaison des dialectes modernes, ont permis de restituer avec une exactitude suffisante la prononciation des caractères chinois à partir du VI[e] siècle. On en trouvera une bibliographie raisonnée dans l'ouvrage fondamental de M. B. Karlgren, *Études sur la phonologie chinoise* (*Archives d'études orientales*, Stockholm, 1915-19). Il convient d'ajouter l'importante étude de M. H. Maspéro, *Le dialecte de Tch'ang-ngan sous les T'ang* (*Bulletin de l'École Française d'Extrême-Orient*, 1920), un article de M. B. Karlgren, *Le proto-chinois, langue flexionnelle*, paru dans le *Journal Asiatique* de 1920 et, du même auteur, *Sound and Symbol in Chinese*, Londres, 1923.

Le chinois moderne, tel qu'on le parle à Pékin et dans les pays de langue mandarine, est caractérisé phonétiquement par l'absence d'occlusives sonores à l'initiale, et de consonnes finales autres que les nasales : un mot ne commence jamais par *b*, *d*, *g*, et il est toujours terminé soit par une voyelle, soit par une des deux nasales *n* ou *ṅ*. Cette pauvreté phonétique est un fait relativement moderne. Au VII[e] siècle de notre ère, le système pho-

nétique du chinois était au contraire extrêmement riche. Outre une série de consonnes mouillées, la langue possédait alors une grande variété de consonnes initiales : sourdes, sourdes aspirées, sonores, spirantes. A la finale, on trouvait, outre des voyelles, trois nasales : *m*, *n*, *ṅ* et autant d'implosives : *p*, *ð*, *k* (la nature spirante de la dentale finale est encore controversée). Le vocalisme était également très riche et comprenait de nombreuses diphtongues et triphtongues. On distinguait enfin quatre tons dont chacun se prononçait à un niveau supérieur quand le mot commençait par une sourde et à un niveau inférieur quand l'initiale était sonore ; il y avait donc en réalité huit tons.

Si l'on réserve pour la période antérieure au VIe siècle ap. J.-C. la dénomination de « chinois archaïque », on peut distinguer trois grandes périodes dans l'évolution phonétique du chinois postérieurement à cette date :

1) la période du chinois ancien s'étend du VIe au Xe siècle. Il se forme alors une série dentilabiale aux dépens des bilabiales et les anciennes occlusives initiales sonores deviennent des sonores aspirées.

2) le chinois moyen dure du Xe au XIIIe siècle : pendant cette période de transition, le système des finales commence à se simplifier dans les dialectes du Nord et les initiales sonores aspirées poursuivent leur évolution ; le traitement de ces dernières varie suivant les dialectes, mais en général elles tendent à se transformer en sourdes aspirées ou non aspirées.

3) au début de la période du chinois moderne, *m* final a disparu et a été remplacé par *n* en chinois du Nord.

Dès le début de la période ancienne, c'est-à-dire au VIe siècle, les groupes initiaux consonantiques étaient presque tous réduits en chinois. Par analogie avec le tibéto-birman, on peut supposer que le chinois archaïque avait un système de préfixes et de suffixes, mais dès le sixième siècle, ces éléments avaient presque complètement disparu ou s'étaient amalgamés avec la racine.

Pendant la période ancienne, la morphologie est déjà très

simple ; la langue est aussi analytique que possible. Le mot est invariable, quel que soit son rôle dans la phrase. Bien mieux, certains mots, sans changer de forme, peuvent être employés comme noms ou comme verbes, comme verbes ou adjectifs, etc. Ces variations sémantiques ne sont rendues intelligibles que par l'emploi de particules et surtout par la place assignée au mot dans la phrase : on comprend par exemple que tel mot est un nom au cas-régime parce qu'il suit immédiatement un verbe ; et tel mot ordinairement employé comme substantif devient verbe dans une phrase où il est suivi d'un régime.

En était-il de même en chinois archaïque ? Il y a des raisons d'en douter. Certains faits semblent indiquer qu'à une époque très reculée, les racines n'étaient point absolument fixes et immuables ; elles étaient susceptibles de diverses transformations suivant le sens qu'on leur donnait et le rôle qu'elles devaient jouer dans la phrase.

Certains caractères apparentés graphiquement et sémantiquement avaient à l'époque ancienne des prononciations similaires mais toutefois un peu différentes. C'est ainsi qu'il existe deux caractères analogues, prononcés tous deux *pe* en langue mandarine, dont l'un signifie « dos » et l'autre, qui désigne le Nord, signifie au propre « tourner le dos », le Nord étant la région qu'on aurait dans le dos en regardant le Midi. Le premier caractère à l'époque ancienne était prononcé en syllabe ouverte, tandis que le second avait une gutturale finale. La même racine avait par conséquent deux formes différentes suivant qu'elle était prise en valeur de substantif « dos » ou en valeur de verbe « tourner le dos ».

Le même caractère signifie tantôt « long », tantôt « croître, grandir ». Dans les deux cas, il est prononcé *tšaṅ* ou *tš'aṅ* mais sur des tons différents : l'adjectif appartient à la série basse et le verbe à la série haute. Il en ressort qu'à l'époque ancienne, le premier avait une initiale sonore et le second une initiale sourde. J'ai signalé plus haut des alternances du même type en tibéto-birman : dans cette famille linguistique, de même qu'en chinois, certaines formes à initiale sonore alternent avec

des formes à initiale sourde, ces dernières ayant généralement une valeur causative ou transitive.

D'autre part, M. B. Karlgren a montré qu'en chinois archaïque les pronoms personnels avaient des formes différentes au cas-sujet et au cas-régime (voir ci-dessus, p. 364).

Pour importants qu'ils soient, ces vestiges d'un passé lointain ne donnent qu'une idée très insuffisante de la morphologie du chinois archaïque. Ils permettent du moins de le caractériser comme une langue aux racines souples et muables, susceptibles d'exprimer des sens divers et des rapports grammaticaux en modifiant leurs éléments et en s'adjoignant des morphèmes, par opposition avec la langue moderne où s'est depuis longtemps perdue toute notion de dérivation, où noms, verbes, adjectifs, lorsqu'ils sont isolés, ne se reconnaissent à aucun signe, et dont les mots sont extrêmement courts et invariables.

Cet abrègement des mots et l'immutabilité qui en est la conséquence ne sont pas particuliers au chinois moderne ; on peut suivre une évolution parallèle dans la plupart des langues d'Extrême-Orient, notamment en tai et en annamite. Mais ce qui est remarquable en chinois, c'est que la tendance au monosyllabisme paraît s'y manifester avec plus de rigueur et plus tôt que partout ailleurs.

Comment une langue où les mots s'abrègent à l'excès, qui ne tolère aucun groupe consonantique et élimine presque toutes les finales, qui conserve un énorme vocabulaire et persiste à vouloir exprimer toutes les nuances, peut-elle néanmoins rester intelligible ? Le chinois permet de répondre à cette question qui se pose d'ailleurs, presque dans les mêmes termes, pour la plupart des langues dites « monosyllabiques ».

Comme un seul de ces mots trop courts risquait d'être incompris ou confondu avec d'autres, on a pris le parti d'en juxtaposer plusieurs, en sorte que, dans la langue moderne, c'est souvent un groupe de plusieurs mots qui exprime ce que nous rendons par un seul dans nos langues occidentales. Il serait trop long de décrire ces combinaisons. Il suffit d'en citer quelques-unes : l'emploi de composés formés de deux mots presque

synonymes, et l'utilisation, à côté du substantif, d'un « classificateur » destiné à indiquer la catégorie à laquelle appartient l'objet désigné ou à évoquer son image en rappelant un de ses traits caractéristiques.

On observe des procédés identiques dans les autres langues d'Extrême-Orient où s'accuse la tendance au monosyllabisme : langues tai, annamite, etc. On aurait tort d'y chercher des arguments pour démontrer la parenté de ces langues. Partout il a fallu rester intelligible alors que les mots trop courts risquaient d'être confondus ; on a donc combiné les monosyllabes suivant les lois universelles de l'association des idées et il n'est pas étonnant que l'application inconsciente des mêmes principes ait donné partout les mêmes résultats.

LES LANGUES TAI

La plus connue des langues tai est le siamois. C'est une langue de civilisation parlée par environ six millions d'individus et qui a donné naissance à une abondante littérature. Les autres parlers sont répandus sur une aire immense depuis la Chine jusqu'à l'Inde. Bien que disséminées et connues, pour la plupart, uniquement sous leur forme la plus moderne, ces langues sont très peu différentes les unes des autres et constituent une famille remarquablement homogène. Les tribus tai sont parfois aussi appelées *tšām* ou *šām*, d'un nom dont la forme birmane est *šān*, et d'où paraît dériver « Siam ». Quant au mot *tai*, c'est une appellation ethnique à ancienne initiale sonore **dai* qui est devenue *t'ai* en siamois et en laotien, c'est-à-dire dans les langues du Sud, et *tai* dans les langues septentrionales. Ce mot illustre ainsi une loi phonétique très importante : les anciennes initiales sonores encore attestées par l'écriture se sont changées en sourdes aspirées dans les langues tai du Sud : siamois et laotien, et en sourdes non aspirées dans la plupart des langues du Nord : parlers chān, khām-tī, tai noir, tai blanc, thô et dioi.

En principe, les individus qui parlent des dialectes appartenant au groupe nord peuvent se comprendre, de même que les individus parlant des dialectes du groupe sud ; mais la différence entre les deux groupes de parlers est plus sensible.

Sur les diverses langues tai on trouvera des indications bibliographiques dans *Linguistic Survey of India*, vol. II, et dans *Bibliotheca Indosinica* de H. Cordier.

Un certain nombre de langues tai : siamois, laotien, āhom,

etc., sont écrites avec des alphabets originaires de l'Inde. L'écriture siamoise fut créée en 1284 de notre ère. La plus ancienne inscription connue en langue siamoise est datée de 1293. Pour l'épigraphie tai, les sources principales sont *La mission Pavie*, Études diverses, II, Paris, 1898, et le *Siam Ancien* de Fournereau, *Annales du musée Guimet*, t. XXXI.

En Birmanie vivent les Karen dont la langue assez mal connue présente certaines affinités avec le chinois et le tai.

D'autre part, dans certains districts du Sud de la Chine et du Nord de l'Indo-Chine, campent des groupes de montagnards appelés Man, Miao-tseu, Meo, etc. Les diverses tribus s'étagent souvent à de grandes hauteurs, sur les pentes des montagnes. Leurs vocabulaires sont chargés de mots tai, chinois, ou annamites suivant la langue parlée par leurs voisins immédiats. La comparaison de ces parlers reste à faire et il est actuellement impossible de dire s'ils se rattachent à la famille tai, au chinois ou à l'annamite, bien que la première de ces alternatives soit assez vraisemblable.

Aux premiers siècles de notre ère, les régions au Sud du fleuve Bleu étaient probablement peuplées de Tai et de Miao-tseu. Plus tard, refoulées par les Chinois, ces populations durent émigrer vers le Sud et elles envahirent l'Indo-Chine. Toutefois de nombreux ilots préchinois se sont maintenus jusqu'à ce jour en Chine depuis le Sseu-tch'ouan et le Yunnan jusque dans l'île de Hainan.

Les tribus tai qui descendirent la vallée du Mékong fondèrent au XIIIe siècle le royaume de Siam qui empiéta progressivement sur l'empire khmer. Un groupe oriental pénétra jusqu'en Assam vers 1228. Ce furent les Āhom. Bien que peu nombreux, ceux-ci réussirent à fonder un petit royaume et à sauvegarder longtemps leur indépendance. Actuellement l'āhom est une langue morte.

Le changement d'habitat des Tai, attesté par les documents historiques, pourrait être invoqué pour expliquer l'assourdissement des initiales sonores qu'on observe dans les langues tai On est tenté de supposer qu'à partir d'une certaine époque, ces

langues auraient été parlées par des populations incapables de prononcer des initiales sonores. Cette conjecture est rendue plus vraisemblable par le fait que les tribus siamoises ont colonisé un territoire déjà conquis partiellement par les Cambodgiens et que le même phénomène d'assourdissement des sonores, limité toutefois à certains phonèmes, s'observe précisément en khmer.

Un autre caractère de la phonétique des langues tai est l'existence en tai commun d'une catégorie spéciale d'occlusives intermédiaires entre les sonores et les sourdes. Ces phonèmes qu'on désigne souvent du nom de mi-sourdes et qu'on peut transcrire ƀ et đ, sont notés dans l'écriture siamoise par des signes spéciaux. Ils se sont maintenus jusqu'à présent dans la plupart des langues tai. Ils ont disparu en chān et en khām-tī. Dans ces deux langues, đ est passé à *l* et ƀ est devenu *m*.

En tai commun, les diverses catégories d'initiales avaient sur la hauteur du ton une influence déterminante. D'une façon analogue à ce qui a déjà été observé en tibéto-birman et en chinois, les mots commençant par une sonore étaient prononcés sur un ton plus bas, les mots commençant par une sourde sur un ton plus élevé. Quant aux mi-sourdes, elles avaient apparemment sur le ton la même influence que les sourdes.

Malgré l'usure qui, dans toutes les langues tai, a beaucoup simplifié la morphologie, l'āhom et le siamois ont conservé un certain nombre de préfixes et apparaissent pour cette raison comme des idiomes relativement archaïques.

Au début des mots, l'āhom admet, devant *l* ou *r*, un assez grand nombre de consonnes : *p*', *k*', *m*, *b*, *p*, *t*, *k*. De même, en siamois, les groupes initiaux consonantiques ont toujours pour second élément une liquide. Il s'en faut d'ailleurs que racine et préfixe se laissent toujours aisément dissocier. En présence, par exemple, d'un groupe initial : *p'r*, on ne peut pas décider *a priori* s'il s'agit d'une ancienne racine à initiale *r* précédée d'un préfixe *p*' ou d'une racine à initiale *p*' suivie de *r*.

La tendance à la réduction des groupes initiaux paraît s'être

manifestée dans toutes les langues tai. Même en āhom où subsistent des groupes nombreux, le système des préfixes s'est fortement altéré et leur rôle ancien est inconnu. Il en résulte que la formation des mots dans les langues tai reste enveloppée d'obscurité.

Les substantifs restent invariables ; leur fonction grammaticale ressort de leur position dans la phrase ou de la présence de certaines particules. Dans la plupart des langues tai, le sujet précède le verbe, et le régime suit le verbe ; mais il arrive souvent en āhom et quelquefois en siamois que le régime précède le verbe.

Les adjectifs sont tantôt des adjectifs proprement dits, tantôt des verbes qui jouent le rôle d'adjectifs ; ces derniers peuvent être alors accompagnés d'une particule qui néanmoins fait défaut dans un grand nombre de cas.

Les verbes sont invariables comme les autres mots. Rien dans la forme ne les distingue d'un substantif. On reconnaît un verbe à sa position dans la phrase et souvent en outre au fait qu'il est accompagné de préverbes, de particules verbales et d'auxiliaires. Mais un mot qui a le sens d'un verbe peut, dans bien des cas, être employé indifféremment comme adjectif, verbe passif ou verbe actif. Il en résulte que la distinction classique entre phrases nominales et verbales perd beaucoup de son intérêt dans les langues tai.

Si un même mot équivaut souvent, sans modification de forme, à un adjectif ou à un verbe, des variations s'observent, dans certains cas, suivant que le verbe est actif ou causatif d'une part, intransitif ou passif d'autre part. L'alternance est du type : verbe intransitif à initiale sonore, verbe transitif à initiale sourde ou sourde aspirée. Ainsi en siamois :

gan « être fondu », *kăn* « séparer »
gĕn « être blessé », *k'ĕn* « opprimer, fouler »

On observe les mêmes alternances entre des formes nominales et verbales. Par exemple, en siamois :

gŏt « rouleau » { *k'ot* « lier, enserrer » / *kot* « embrasser » }
bu « abcès » *p'u* « putréfier »

Ces variations, parallèles à celles qu'on observe en chinois et en tibéto-birman, sont de nature à faire classer les langues tai dans le groupe sino-tibétain. De même, la relation entre la hauteur du ton et la nature de l'initiale. Il est à noter également que les noms de nombre, dans l'ensemble, sont les mêmes en chinois, en tibéto-birman et dans les langues tai. Aucun de ces rapprochements n'a d'ailleurs une importance décisive. Il est possible que la prononciation des mêmes phonèmes initiaux détermine automatiquement la même accentuation dans des langues non apparentées. Quant aux ressemblances dans la série des nombres et même dans la forme des mots, elles s'expliquent peut-être par d'anciens emprunts au groupe sino-tibétain.

Par contre, les formes pronominales tai présentent de curieuses analogies avec celles des langues austroasiatiques.

Le pronom personnel de la première personne est *ao* dans diverses langues mon-khmer, *tāo* en annamite, *kāo, kāu* dans les langues tai. On verra plus loin que, dans les langues mon-khmer, les préfixes *ka, k,* ou *ta, t,* se placent indifféremment devant les racines qui désignent des êtres animés. Ainsi pourraient s'expliquer, à partir du mon-khmer, les formes tai et annamite. Le pronom personnel de la deuxième personne présente également des formes comparables en mon-khmer, en tai et en annamite.

Dans l'emploi des démonstratifs, le tai, ainsi qu'un grand nombre de langues austroasiatiques, distingue trois degrés suivant que l'objet désigné est rapproché, plus éloigné ou encore plus éloigné.

	Siamois	Annamite	Khasi	Palaung
Position rapprochée.......	*ni*	*nay, dey*[1]	*ne*	*ö*
Position plus éloignée.....	*nan*	*(n)ey, dey*[1]	*to*	*nān*
Position encore plus éloignée	*nōn*	*no, do*	*tai*	*tāi*

1. *dey* signifiant « là » se prononce sur un ton élevé, tandis que *dey* signifiant « ici » se prononce sur un ton moyen. De même, *(n)ey* est prononcé sur un ton plus élevé que *nay*.

Ce qui est remarquable ici, au moins autant que la ressemblance des mots d'une langue à l'autre, c'est que, dans un même idiome, les divers degrés d'éloignement s'expriment souvent par des variations d'une même racine : *nï*, *nan*, *nōn*, etc.

Au total, les langues tai ont des analogies avec le sino-tibétain et par d'autres côtés elles ressemblent aux langues mon-khmer. C'est pourquoi, bien que la plupart des auteurs les rattachent au premier groupe, on peut être tenté de les classer parmi les langues austroasiatiques. Tout s'expliquerait aisément si l'on prouvait que le sino-tibétain et les langues austroasiatiques appartiennent à une même famille linguistique et que les langues tai sont apparentées à la fois à l'un et aux autres. Cette thèse a été soutenue récemment par M. Conrady dans *Aufsätze zur Kultur- und Sprachgeschichte vornehmlich des Orients, Ernst Kuhn gewidmet...*, p. 475-504. On ne peut dire que son argumentation soit absolument convaincante. Dans l'état actuel de nos connaissances, les chances d'erreur restent nécessairement nombreuses, et quand on croit apercevoir la vérité il n'est guère possible d'en donner une démonstration rigoureuse.

J. Przyluski.

LES LANGUES AUSTROASIATIQUES

GÉNÉRALITÉS

L'expression « langues austroasiatiques » a été proposée par le P. W. Schmidt pour désigner un groupe de langues parlées depuis l'Annam à l'Est jusqu'au plateau de Chota Nagpour à l'ouest (voir planche 11). Le vocabulaire et la grammaire de ces langues présentent certaines analogies, et le P. Schmidt admet que tous ces idiomes appartiennent à une même famille linguistique. Celle-ci ne serait elle-même qu'une partie d'un ensemble encore plus vaste, comprenant la totalité des langues malayo-polynésiennes, et que le savant linguiste appelle « famille des langues austriennes ». La construction est grandiose ; elle est encore assez fragile. Il se peut que les recherches ultérieures confirment les idées du P. Schmidt. Il convient actuellement de ne les point admettre sans réserve.

Il est probable que les langues austroasiatiques se sont superposées à des langues plus anciennes dont il reste çà et là des débris, notamment chez les tribus de la Péninsule malaise. D'autre part, pendant les derniers millénaires, des influences diverses : chinoise, malaise, tibéto-birmane, aryenne, etc., se sont exercées sur les langues austroasiatiques et les ont ici refoulées, là recouvertes d'apports nouveaux. Entre le substrat ancien et les sédiments récents, la couche austroasiatique est parfois très mince, et son extension réelle ne peut être évaluée que grâce à certains indices dont l'interprétation est, en bien des points, conjecturale.

Dans ces conditions, la théorie des langues austroasiatiques

apparaît comme une hypothèse destinée à rendre compte de certaines analogies et qui, tout en ayant un caractère incertain et provisoire, permet d'expliquer, de la manière la plus probable, la plupart des faits connus à ce jour.

Si, au lieu de considérer l'ensemble des langues austroasiatiques, on se borne à examiner des parties de ce vaste domaine, on y observe plusieurs groupes linguistiques moins étendus dont la cohésion et l'unité sont beaucoup plus sûres : à l'Ouest les langues muṇḍā, ou kōl ; à l'Est l'annamite ; au Centre le vaste groupe mon-khmer dont les principaux idiomes sont sans doute apparentés, mais qui comprend en outre une foule de parlers dont un grand nombre sont mal connus et ne se laissent par conséquent déterminer qu'avec de fortes chances d'erreur. Nous nous attacherons principalement à dégager les caractères généraux de ces divers groupes.

En ce qui concerne la totalité des langues austroasiatiques considérées comme faisant partie d'un même ensemble linguistique, le mieux qu'on puisse faire actuellement est d'énumérer les analogies que présentent les trois groupes : mon-khmer, annamite et muṇḍā (kōl).

Les premiers linguistes paraissent avoir été guidés principalement dans la classification de ces langues par la diffusion des noms de nombre. A cet égard, les trois groupes austroasiatiques sont remarquablement voisins comme suffit à l'indiquer la comparaison des premiers nombres en santālī (muṇḍā), en mon (mon-khmer) et en muong (annamite) :

	santālī	mon	muong
un	*mit*	*mwai*	*môt*
deux	*bar*	*bă*	*hal*
trois	*pä*	*pi*	*pa*
quatre	*pōn*	*pan*	*pôn*

Plus convaincantes sont les analogies dans les procédés qui servent à former les mots dérivés. C'est ainsi que l'infixe *n* paraît jouer le même rôle dans les trois groupes austroasiatiques

où il forme généralement des noms d'instrument et des adjectifs ou des mots abstraits. Cette similitude est d'autant plus remarquable que les anciens procédés de dérivation sont encore très peu connus.

Dans l'état de décomposition où sont la plupart des langues modernes, la morphologie du nom et du verbe est le plus souvent réduite à une extrême simplicité ; mais les pronoms paraissent avoir mieux conservé les formes anciennes. On a déjà signalé, à propos des langues tai, que les démonstratifs annamites, mon-khmer et muṇḍā ont au moins trois formes suivant l'éloignement de l'objet désigné et que, pour une même langue, la majorité sinon la totalité de ces formes résultent souvent du développement d'une même racine. D'autre part les pronoms personnels ont des formes pour le duel dans l'ensemble des parlers muṇḍā et dans quelques-unes des langues mon-khmer. On observe en outre sur tout le domaine austroasiatique des formes pronominales « inclusives » et « exclusives » employées suivant que le sujet parlant comprend ou non son interlocuteur parmi les personnes qu'il désigne.

Enfin, d'après le P. Schmidt, « la caractéristique la plus intime » des langues austroasiatiques serait la relation constante entre certaines finales et des groupes sémantiques déterminés. Par exemple, les finales en *p* se présenteraient constamment dans des mots qui expriment l'idée d' « étreindre, saisir, pincer, tenailler », puis de « mordre, manger, mâcher », et aussi en général de « goûter » puis de « boire », puis de « couvrir » et encore l'idée générale d' « ensemble », « se rencontrer, se contracter, s'atrophier » (*Les peuples mon-khmer, trait d'union entre les peuples de l'Asie Centrale et de l'Austronésie*, trad. franç. dans *Bulletin de l'École Française d'Extrême-Orient*, 1907, p. 243). D'après le même auteur, cette relation entre telle finale et tel groupe sémantique serait due à d'anciens suffixes qui se seraient soudés à la racine. Il y a là une indication fort importante susceptible d'orienter les recherches dans une direction nouvelle.

La parenté des langues austroasiatiques, si elle était démontrée, serait de grande conséquence pour l'historien du langage, car ainsi seraient rapprochés des idiomes qu'on avait accoutumé de classer dans des catégories toutes différentes : d'une part le muṇḍā qui présente tous les caractères des langues dites « agglutinantes » et d'autre part l'annamite éminemment « monosyllabique ». Il ne faut pas d'ailleurs s'exagérer l'intervalle qui sépare ces deux groupes et on conçoit qu'une même langue ait pu donner naissance à l'annamite et au muṇḍā. Voici en effet comment, à défaut de textes anciens, on pourrait se représenter l'évolution des langues austroasiatiques à partir d'un ancêtre commun.

On verra plus loin que certains caractères communs à l'annamite, au mon-khmer et au muṇḍā permettent d'imaginer un état ancien de ces langues où la liaison des mots dans la phrase était suffisamment assurée par les morphèmes agglomérés à la racine. Dans les langues modernes au contraire, et principalement dans celles de l'Est, le système de dérivation, la morphologie sont en décadence ; il en résulte que les éléments de la phrase n'ont plus assez de cohésion et ne peuvent marquer par eux-mêmes les rapports qui les unissent. Pour conserver au discours son unité et sa continuité, il a fallu recourir à des procédés qui varient beaucoup suivant les langues considérées. Dans les parlers orientaux, tels que l'annamite, on a coulé entre les mots ce ciment que sont les « particules » et on a pris l'habitude d'assembler, par la diction comme par la pensée, un certain nombre de mots destinés à former un « groupe » où chaque élément a son rôle et sa valeur fixés d'avance. La syntaxe, aidée du rythme, s'est substituée à la morphologie défaillante. A l'Ouest la tâche était plus aisée parce que les morphèmes s'étaient maintenus en assez grand nombre ; toutefois, leur valeur s'étant obscurcie, on a renforcé la cohésion de la phrase en multipliant les pronoms qui, de tous les éléments du discours, sont restés les plus souples et les plus variés. Partout, en somme, le mot devenant impropre à cette double fonction : exprimer un sens précis et marquer des rapports grammaticaux, on a con-

fié ce rôle à des « groupes », à des formations syntaxiques intermédiaires entre la phrase et le mot.

Considérons par exemple les deux « groupes » : santālī *em-äe-mä* et annamite *cho nó đi* ; tous deux signifient : « donne-lui. . . » et sont formés le premier d'un élément verbal *em* et de deux formes pronominales *äe*, *mä*, le second du verbe *cho* . .*đi* encadrant le pronom *nó*. Entre ces deux formations, la différence est plus apparente que réelle. La première, transcrite par des auteurs habitués aux alphabets indiens, a été notée comme un seul mot de trois syllabes. La seconde, écrite traditionnellement à la chinoise, apparaît comme une suite de trois mots monosyllabiques. Dans l'un et l'autre cas, il s'agit d'un agrégat complexe qui, de quelque manière qu'on l'écrive, déborde les limites du « mot » proprement dit.

LES LANGUES MON-KHMER

Outre un grand nombre d'idiomes parlés par des tribus non civilisées disséminées depuis la chaîne annamitique jusqu'à l'Assam, la famille mon-khmer comprend trois langues de civilisation attestées à date ancienne : le mon, le khmer et le tcham. Sur les inscriptions en vieux-mon, l'*Epigraphia Birmanica* éditée à Rangoon par MM. Taw Sein Ko et Duroiselle, contient les données les plus accessibles et les plus récentes. La plus célèbre de ces inscriptions est celle qui couvre une des faces du pilier de Myazedi (1084 ap. J.-C.?). Les plus anciennes inscriptions khmer connues à ce jour sont de la première moitié du VII[e] siècle. Le vieux-tcham est encore très mal connu. Sur l'épigraphie khmer et tcham, on trouvera une bibliographie dans l'ouvrage de M. Finot, *Notes d'Épigraphie* (*Bulletin de l'École Française d'Extrême-Orient*, 1915, p. 137 et suiv.).

Le deuxième volume de *Pagan Races of the Malay Peninsula* par Skeat et Blagden (Londres, 1906) contient un admirable vocabulaire comparatif et une bibliographie (Part IV et appendice). *Bibliotheca Indosinica* de H. Cordier donne une abondante bibliographie.

On peut classer géographiquement les langues mon-khmer de la façon suivante :

1) un groupe central comprenant le mon ou talaing, le khmer ou cambodgien, le bahnar, le stieng, le röngao et un grand nombre d'autres parlers « sauvages » : moï de la chaîne annamitique, kha, koui, chong, pear, penong du bassin du Mékong.

2) un groupe oriental comprenant le tcham, le jaraï, le radé et le sedang. Ce groupe a subi profondément l'influence de la langue malaise.

3) le groupe de la péninsule de Malakka comprenant les parlers semang, sakai et jakun.

4) le nikobarais.

5) le groupe du bassin moyen de la Salouen comprenant le palaung, le wa et le riang et auquel on peut rattacher le khamûk ou khmu et le le-met.

6) le khasi parlé en Assam par 177.000 individus isolés au milieu d'une population linguistiquement différente.

Cette classification est principalement géographique et on ne saurait lui attribuer *a priori* une valeur linguistique. Tel parler de la péninsule de Malakka, par exemple, peut avoir plus d'affinités avec une langue du groupe central qu'avec le parler d'une tribu voisine.

En présence de langues aussi disséminées, dont la plupart n'ont jamais été fixées par l'écriture et qui ont évolué dans des circonstances toutes différentes, on pourrait s'attendre à ne plus trouver aujourd'hui que des formes disparates et très divergentes. C'est le contraire qu'on constate avec surprise. Des mots, des procédés de dérivation reparaissent presque identiques dans de nombreux idiomes de cette famille.

La phonétique des langues mon-khmer est remarquablement stable, au moins dans certains de ses éléments. Cette stabilité est d'ailleurs plus sensible chez les « sauvages » que chez les civilisés. C'est ainsi que les Cambodgiens, fondateurs d'un empire, et qui ont été longtemps en contact avec les peuples voisins, parlent une langue phonétiquement plus évoluée que celle des tribus qui vivent à l'écart dans les forêts et les montagnes. A cet égard l'écriture cambodgienne permet de mesurer jusqu'à un certain point le chemin parcouru car elle conserve, sans doute assez fidèlement, la forme ancienne des mots, tandis que la prononciation moderne est sensiblement différente.

« montagne »	écrit	*b'nam*	est prononcé	*p'nom*
« bois »	«	*dẓ̌'ö*	«	*tš'ö*
« loi »	«	*d'arm(a)*	«	*t'or*
« ville »	«	*nagar(a)*	«	*nokor*

Ces exemples laissent apercevoir que les anciennes occlusives sonores sont devenues sourdes (la sonore *ḍ* toutefois fait exception). Inversement, les deux sourdes *ṭ* et *p* sont devenues sonores.

Des tendances analogues se sont manifestées dans plusieurs langues mon-khmer, notamment en mon, autre langue de civilisation. Toutefois c'est en khmer qu'elles paraissent s'être réalisées avec le plus d'ampleur.

Encore que les mots khmer soient relativement usés, la manière dont ils ont été formés se décèle encore aujourd'hui, en sorte qu'on peut grouper les dérivés autour de la racine qui leur a donné naissance. Ex :

kăt « couper »
k'năt « mesure »
kŏmnăt « morceau »
t'kăt « maladie »
dâmkăt « maladie »
skăt « barrer, couper »
săñkăt « division »
bâñkăt « trancher, décider »

Des faits analogues s'observent aussi nettement dans l'ensemble de la famille mon-khmer.

Des trois grands procédés de dérivation : préfixation, infixation, suffixation, les deux premiers seuls sont actuellement en usage dans la plupart des langues de cette famille. Le nikobarais fait exception : les dérivés par suffixation y sont fréquents. Comme il s'agit d'une langue excentrique par sa position géographique et que le même caractère s'observe dans les langues muṇḍā, il est difficile de décider si le nikobarais a innové sous des influences occidentales ou s'il a seul conservé un trait ancien, aboli depuis longtemps dans les langues parlées à l'est.

On distingue en mon-khmer deux degrés de préfixation : a) un seul préfixe est placé en contact avec la racine ; b) la racine est précédée de deux préfixes. Un préfixe peut lui-même être simple ou composé.

Le préfixe simple, c'est-à-dire ne comprenant qu'une seule

consonne, peut se présenter sous trois aspects : 1) une forme syllabique du type *ka*, *ki*, etc. ; 2) une forme intermédiaire qu'on peut noter *kə*, *ə* représentant un son vocalique très peu distinct ; 3) une forme asyllabique du type *k*. Dans bien des cas, ces trois formes s'observent concurremment dans les diverses langues mon-khmer ; il arrive aussi que le préfixe ait disparu sans laisser de traces. Il semble alors que le mot se soit progressivement réduit suivant un procès dont les principales phases seraient :

ka . . . kə . . . k . . . zéro . . .

Ex. : « aigle de mer » nikobarais *kalañ* ; khmer *k'leñ* ; stieng *kleñ*.
« buffle » semang *kibao* ; jaraï *kəbau* ; stieng *kbau*.
« neuf » mon *tami* ; khmer *t'mīg* ; stieng *mëi*.

Les préfixes simples sont très nombreux. Leur rôle est difficile à déterminer. Les préfixes *ka* et *ta* ont servi sans doute à former des substantifs, surtout des noms d'animaux, de plantes, de parties du corps. Quant au préfixe *pa*, il formait des verbes causatifs. Ce dernier point est un des mieux établis de la grammaire comparée des langues mon-khmer. Exemple :

vieux-mon *ār* « aller » *pār* « faire aller », « se conduire ».

Les préfixes composés, c'est-à-dire comprenant deux consonnes, sont formés théoriquement d'un préfixe simple auquel se serait ajoutée une nasale ou une liquide. On a ainsi des groupes très variés : *kan*, *kar*, *kal*, *kəm*, etc.. Leur rôle est assez mal défini.

Les cas de préfixation du deuxième degré sont assez rares. Le palaung en offre des exemples très nets :

yām « mourir », *p-yām* « tuer », *pan-p-yām* « mise à mort ».

L'infixation consiste dans l'insertion d'un élément simple ou composé à l'intérieur de la racine. Les infixes simples sont généralement des liquides ou des nasales ; on rencontre aussi un infixe *p*. L'infixe composé *mn* n'a été trouvé qu'en khmer et en nikobarais et *rl* uniquement en khmer. L'infixe *n*, qui est de beaucoup le plus fréquent, forme généralement dans les langues mon-khmer des noms d'instrument et des adjectifs. Ex. :

mon *put* « ciseler », *pnut* « ciseau »
khasi *kiau* « grand'mère », *kəniau* « vieux »

Une racine mon-khmer étant susceptible de donner naissance à des substantifs, verbes et adjectifs, par l'adjonction de préfixes et d'infixes appropriés, il en résulte que ces catégories de mots avaient jadis certains caractères distinctifs qui empêchaient de les confondre. Dans les langues modernes, la réduction de la longueur des mots a presque partout ruiné l'ancien système en sorte que le sujet parlant n'a plus guère conscience du rôle des préfixes et des infixes et qu'il est devenu souvent malaisé, parfois même impossible, de distinguer un verbe d'un substantif ou d'un adjectif, indépendamment de leur sens ou de leur emploi dans une phrase donnée.

En principe, le mot est invariable, quelle que soit sa fonction grammaticale ; les rapports des noms entre eux, les circonstances de genre et de nombre, les temps et modes des verbes, tout ceci est rendu au moyen de particules supplémentaires et par la place qu'occupe le mot dans la phrase. Sauf exception, l'ordre des mots est : sujet, verbe, régime.

Néanmoins, divers indices donnent à penser qu'il y eut un temps où certaines de ces nuances s'exprimaient par des modifications de la racine. A cet égard, les pronoms personnels du khasi sont particulièrement instructifs :

	Masculin	Féminin	Pluriel
1re personne..	*ṅā*	*ṅā*	*ṅi*
2e personne..	*mē*	*p'ā*	*p'i*
3e personne	*u*	*ka*	*ki*

L'alternance ā/i du singulier au pluriel est extrêmement remarquable dans une langue où substantifs et verbes ne présentent plus rien de pareil. Dans le même ordre d'idées, j'ai déjà signalé plus haut, à propos des langues tai, les variations des démonstratifs mon-khmer suivant que l'objet désigné est proche, éloigné ou plus éloigné (voir ci-dessus, p. 383). Il est à noter également qu'en palaung et en nikobarais les pronoms personnels ont une forme de duel. Ainsi, dans les langues mon-khmer, comme en sino-tibétain, le pronom est, parmi les éléments du discours, un de ceux qui paraissent avoir conservé le plus longtemps les anciennes formes grammaticales.

L'ANNAMITE

La famille annamite comprend une langue de civilisation : l'annamite et un certain nombre de parlers en usage chez les Muong du Tonkin et de l'Annam. L'ensemble est remarquablement homogène. On peut évaluer à 14 millions le nombre des Annamites.

Une étude sérieuse de ce groupe linguistique devrait reposer sur la comparaison de l'annamite et des parlers muong ; malheureusement ces derniers ne sont connus que par des vocabulaires assez pauvres.

Les Annamites qui n'écrivaient au début qu'en chinois, inventèrent plus tard pour transcrire leur langue un système de caractères formés d'éléments empruntés à l'écriture chinoise : les *chữ-nôm* ou caractères vulgaires. Une inscription prouve que ces *chữ-nôm* existaient dès le XIV[e] siècle, mais le plus ancien livre en caractères annamites qui nous soit parvenu est du XV[e] siècle. Au XVII[e] siècle, les missionnaires européens écrivirent l'annamite avec des caractères latins. Ce fut le *qu'ôc-ngữ*. Depuis lors, de nombreux ouvrages ont été publiés avec ce mode de transcription.

On trouvera une liste d'ouvrages relatifs à l'annamite dans *Bibliotheca Indosinica* de H. Cordier.

Les premiers linguistes qui étudièrent l'annamite le classèrent avec les langues mon-khmer et créèrent une famille : mon-khmer-annamite. Puis on a réagi contre cette tendance et préféré rattacher l'annamite aux langues septentrionales : chinois et tai. Il semble que les premiers linguistes avaient raison et qu'il soit temps de leur rendre justice.

L'annamite et les parlers muong ont un système de tons tandis que les langues mon-khmer sont toutes, autant qu'on sache, monotones. Cette différence est un des principaux arguments de ceux qui nient la parenté de l'annamite et du mon-khmer. Mais aussi longtemps qu'on ignorera les circonstances où une langue donnée perd ou conserve un système de tons, il sera prudent de ne pas faire état de la disparition ou du maintien d'un tel système, lorsqu'on voudra déterminer la généalogie des langues.

Le vocabulaire annamite se compose en majeure partie de mots de civilisation empruntés au chinois et, pour une moindre part, de mots très usuels dont un bon nombre se retrouvent encore aujourd'hui en mon-khmer. Cette constatation est en faveur du rattachement de l'annamite aux langues austroasiatiques plutôt qu'au sino-tibétain.

On a vu précédemment que les anciennes sourdes *ṭ* et *p* sont devenues sonores en khmer après la création de l'écriture cambodgienne. Le même phénomène s'est produit parallèlement en annamite, tandis que les parlers muong ont généralement conservé la sourde.

	Khmer classique	Mon	Muöng	Khmer moderne	Annamite
trois	*piy*	*pi*	*pa*	*bey*	*ba*
quatre	*puon*	*pan*	*pôn*	*buon*	*bôn*
terre	*ṭiy*	*ti*	*tät*	*dey*	*dät*
nuit		*betâm*	*têm*		*dêm*

L'ancienne gutturale sourde *k* s'est maintenue sans modification en annamite comme dans la plupart des langues mon-khmer.

	Mon	Stieng	Bahnar	Annamite
enfant	*kōn*	*kôn*	*kon*	*kon*
poisson	*ka*	*ka*	*ka*	*ka*

Dans bien des cas, les correspondances phonétiques entre l'annamite et le mon-khmer sont moins évidentes que dans les exemples cités plus haut, parce que l'usure a souvent défiguré les mots du premier groupe. En annamite moderne, en effet, la

longueur des mots est toujours réduite au strict minimum. Ceci est particulièrement apparent lorsqu'on considère l'évolution des préfixes.

Quiconque se bornerait à observer l'annamite tel qu'on le parle aujourd'hui aurait quelque peine à deviner que cette langue ait jamais possédé des préfixes. Mais il suffit de remonter au XVII[e] siècle pour en trouver quelques vestiges. Le P. A. de Rhodes, dans son dictionnaire publié en 1649, note trois groupes consonantiques initiaux : *bl*, *ml*, *tl* qui se sont réduits depuis cette époque. La disparition des préfixes, achevée en annamite moderne, n'est pas encore accomplie dans les parlers voisins. Quelques dialectes muong conservent encore aujourd'hui trois groupes initiaux *kl*, *pl*, *tl*. Il ressort de la comparaison de l'annamite et du muong que la langue ancienne avait au moins quatre préfixes *k*, *t*, *p*, *m*. Au XVII[e] siècle, ce système était déjà très altéré puisque les préfixes n'étaient maintenus que devant les racines commençant par *l*.

La réduction des groupes consonantiques initiaux n'a pas eu seulement pour résultat d'amalgamer les préfixes à la racine ; elle a fait disparaître également tout vestige des anciens procédés de dérivation par infixation. Un exemple suffit à montrer que l'infixation a pu jouer autrefois en annamite le même rôle que dans les langues mon-khmer. On a vu plus haut que, dans ce dernier groupe, l'infixe *n* sert notamment à former des adjectifs. Il est possible qu'en annamite *kon* « enfant » ait donné par infixation de *n* l'adjectif **knon* devenu *non* « jeune, tendre » après réduction du groupe *kn*.

L'usure considérable qu'ont subie les mots annamites les a complètement défigurés au point qu'il est aujourd'hui très malaisé de grouper les dérivés autour d'une même racine et que les sujets parlants ont depuis longtemps perdu conscience des anciens procédés de dérivation. A cet égard, l'annamite est, parmi les langues austroasiatiques, comparable au chinois dans le groupe sino-tibétain. La morphologie de l'annamite est, pour les mêmes raisons, aussi peu connue que celle du chinois archaïque ou du tai commun ; et l'annamite moderne a nécessairement tous les

caractères des langues où les mots sont devenus extrêmement courts et invariables : absence de déclinaison, de conjugaison et de formes distinctives du nom, de l'adjectif ou du verbe. Les rapports des éléments du discours sont marqués par des particules indépendantes et par la position des mots dans la phrase. En annamite moderne, comme en chinois, comme dans les langues tai, la syntaxe a pris une énorme importance tandis que la morphologie est très simple. En raison de l'influence profonde exercée par la littérature chinoise, la syntaxe annamite est pleine de tours, d'idiotismes et de procédés calqués sur le chinois classique.

Les parties du discours se succèdent en principe dans l'ordre suivant : sujet, verbe, régime. A l'inverse de ce qui a lieu en chinois, le déterminatif suit le nom qu'il détermine. La différence entre la phrase nominale et la phrase verbale n'est guère apparente parce que le même mot, sans changer de forme, peut être verbe ou adjectif : *giời sáng* « il fait jour » signifie littéralement « le ciel luit » aussi bien que « le ciel (est) lumineux ». La phrase interrogative se construit en général, comme en chinois, avec une négation finale. Pour demander : « est-il venu ? », on dit : « il est venu ou non ? » Comme en chinois, la phrase interrogative a souvent un sens négatif : « est-ce que je sais ? » signifie « je ne sais pas ».

Les pronoms sont les seules parties du discours dont la morphologie soit nettement archaïque. On a déjà vu que les démonstratifs notent divers degrés d'éloignement par des variations de ton et des alternances vocaliques (voir ci-dessus, p. 383). Les pronoms personnels qui, comme les démonstratifs, ont de grandes analogies avec ceux des langues mon-khmer, sont diversement nuancés suivant le rang social des interlocuteurs. A la première personne du pluriel, on emploie deux formes différentes suivant que celui à qui l'on parle est ou n'est pas compris parmi les personnes que l'on désigne. De telles formes pronominales « inclusives » et « exclusives » sont un élément caractéristique de la morphologie des langues muṇḍā et s'observent également dans les langues mon-khmer, notamment en bahnar, en palaung et nikobarais.

LES LANGUES MUNDA

Les populations qui parlent les langues muṇḍā ou kōl habitent deux zones complètement séparées par des populations de langues différentes : au Nord, le versant méridional de la chaine himalayenne, au Sud, le plateau de Chota-Nagpour et les districts voisins. Le premier groupe, de beaucoup le moins important, et qu'on peut appeler « himalayen », ne compte guère plus de cent mille individus. Le second groupe, celui du Chota-Nagpour, compte environ trois millions d'individus.

Le groupe septentrional ou himalayen comprend de l'Ouest à l'Est les parlers suivants : manchāṭī ou patan, bunān, ranglōi, kanāshī, kanāwrī, rangkas ou saukiyā, dārmiyā, byāngsī, chaudāngsī, vāyu, khambu, yākhā, limbu, thāmi et dhīmāl.

Le groupe méridional ou du Chota-Nagpour peut se diviser en deux sous-groupes : l'un plus oriental, le sous-groupe kherwārī, avec les langues ou dialectes santālī, muṇḍārī, bhumij, bīrhāṛ, kōḍā, ho, tūrī, asurī, korwā ; l'autre occidental avec les langues kūrkū, khaṛiā, juāng et peut-être aussi le savara et le gadabā.

Il est probable qu'à date ancienne tout le Nord-Est de l'Inde depuis l'Himalaya jusqu'au Golfe du Bengale était couvert de populations parlant des langues muṇḍā. Mais ces aborigènes ont été refoulés sous la poussée des tribus tibéto-birmanes au Nord et à l'Est, des Aryens à l'Ouest et des Dravidiens au Sud. Actuellement, les langues muṇḍā ne sont guère parlées par des masses compactes que sur le plateau de Chota-Nagpour ; partout ailleurs on ne rencontre plus que des îlots qui s'effritent et se désagrègent peu à peu.

Ces langues ne nous sont connues que sous leur forme la plus moderne. Max Müller qui, le premier, les a opposées aux langues dravidiennes a réuni sous le nom de « famille muṇḍā » les parlers du groupe de Chota-Nagpour. Les mêmes langues ont été appelées tour à tour « kōl », « kolariennes » et « kherwariennes ». Suivant l'usage le-plus répandu, nous réservons ce dernier terme pour désigner un sous-groupe et nous conservons à l'ensemble de la famille le nom que lui a donné Max Müller bien que le mot « muṇḍā » désigne spécialement dans l'Inde les Muṇḍā du district de Ranchi ou encore les individus qui parlent le muṇḍārī.

Jusqu'ici les parlers himalayens ont toujours été tenus pour différents des langues muṇḍā. Le Dr Sten Konow, qui a mis au point pour le *Linguistic Survey of India* les monographies relatives aux uns et aux autres, a bien vu les analogies que présentent ces deux séries de langues ; il a néanmoins rattaché les parlers himalayens à la famille tibéto-birmane. Bien que saturés de mots indo-aryens et tibéto-birmans, certains idiomes en usage sur le versant sud de l'Himalaya s'apparentent aux langues muṇḍā par leur structure morphologique et il nous paraît préférable de les classer avec ces dernières. Toutefois, il convient d'observer que sur bien des points, faute de monographies suffisamment précises, la frontière entre les familles tibéto-birmane et muṇḍā ne peut être tracée avec exactitude, et que l'attribution de certains parlers à tel groupe linguistique est nécessairement douteuse et provisoire.

On trouvera des indications bibliographiques et de nombreuses monographies dans le *Linguistic Survey of India*, vol. III, Part I, p. 273 à 567, et vol. IV, p. 1 à 175.

Un trait remarquable de la phonétique des langues muṇḍā est l'existence de phonèmes appelés généralement « semi-consonnes », qui sont analogues aux implosives finales, mais en diffèrent parce qu'ils sont imparfaitement articulés et peu distincts. Ce caractère, commun aux langues du groupe du Chota-Nagpour, s'observe également dans les parlers du groupe himala-

yen et, en dehors de la famille muṇḍā, dans certaines langues mon-khmer : sakai de Malakka, tcham, etc.

Certaines langues muṇḍā, et particulièrement le santālī, présentent des phénomènes d'assimilation vocalique comparables à ceux qu'on observe dans les langues turco-tartares. Ainsi en santālī, le démonstratif *än* devient *en* lorsqu'il est suivi de *a* ou de *o* et il prend la forme *in* lorsqu'il est suivi de *i*.

Les mots muṇḍā dérivés sont formés par redoublement de tout ou partie de la racine ou par l'adjonction d'éléments étrangers à la racine. Tandis que les langues mon-khmer, le nikobarais excepté, ignorent l'emploi des suffixes, les langues muṇḍā forment des dérivés au moyen de préfixes, infixes et suffixes. L'emploi des infixes est surtout fréquent et leur fonction se laisse assez bien reconnaître : l'infixe *n* forme des noms d'instrument, des adjectifs et des noms abstraits et l'infixe *p* donne naissance soit à des verbes réciproques soit à des noms abstraits. On a par exemple en santālī :

dapal « couvrir » et *danapal* « couverture »
dal « frapper » et *dapal* « se frapper mutuellement ».

Les suffixes servent principalement à modifier les formes verbales, de manière à produire des passifs, des réfléchis, des intransitifs, des verbes d'état. Le suffixe *a* qui forme en nikobarais des intransitifs et des passifs, correspond aux suffixes *ok*, *ak* du santālī, *o* du muṇḍārī et *u* du kūrkū qui forment des intransitifs, des passifs et des réfléchis.

L'existence de nombreux affixes et infixes susceptibles de modifier la valeur sémantique des racines indique assez clairement qu'en muṇḍā commun un verbe transitif ou intransitif, un adjectif, un nom se reconnaissaient probablement à certains signes. Il n'en est plus de même aujourd'hui que dans une très faible mesure : dans les langues modernes, on a tendance à prendre indifféremment un même mot en valeur de verbe, de substantif, etc.

La morphologie n'est pas aussi simple que dans les autres langues austroasiatiques. Les catégories grammaticales : nombre, cas, etc., s'expriment souvent, non plus comme en mon-khmer

ou en annamite au moyen de « particules », c'est-à-dire de mots à part, vidés plus ou moins de leur sens et réduits à une sorte de vassalité vis-à-vis des mots principaux, mais au moyen de morphèmes qui ne peuvent exister à part, se juxtaposent et se soudent aux mots et présentent parfois certaines alternances. Ainsi en santālī le duel des substantifs se forme en y ajoutant *-kīn* et le pluriel en ajoutant *-kō*.

La plupart des parlers muṇḍā distinguent trois nombres : singulier, duel, pluriel. On observe le même caractère dans plusieurs langues mon-khmer (voir ci-dessus, p. 394).

Les langues muṇḍā partagent les noms en deux genres suivant qu'ils désignent des êtres animés ou inanimés. La même division se retrouve dans les langues mon-khmer et en annamite.

Le trait caractéristique de la morphologie muṇḍā est la variété des formes pronominales et verbales.

En santālī, le pronom de la première personne varie au singulier, au duel et au pluriel et chacun des deux derniers nombres présente une forme « inclusive » et une forme « exclusive». De plus chacune de ces formes se modifie par le jeu d'alternances ou l'emploi d'affixes déterminés suivant que le pronom est au cas-sujet, au cas-régime ou au génitif.

Dans les parlers du groupe himalayen, les pronoms personnels prennent également des aspects très divers. C'est pourquoi les auteurs anglais classent souvent ces idiomes sous la rubrique « Complex pronominalized languages ».

Les démonstratifs sont également très nuancés. Comme dans les langues mon-khmer et tai, on distingue plusieurs degrés suivant que l'objet désigné est proche, éloigné ou plus éloigné. Le santālī emploie même des formes démonstratives spéciales suivant que l'objet occupe par rapport au sujet une position normale ou latérale, suivant qu'il s'agit d'une chose vue ou entendue, ou encore d'un objet animé ou inanimé.

Le verbe muṇḍā est souvent formé d'une racine redoublée en totalité ou en partie ou modifiée par l'adjonction d'affixes et d'infixes. Indépendamment de ces procédés qui donnent au verbe un sens passif, causatif, réciproque, moyen, etc., divers morphèmes

sont en outre usités pour marquer le temps et le mode. On obtient ainsi, au moins dans certains parlers tels que le santālī, un très grand nombre de formes et il arrive qu'un verbe seul présente une longue suite de syllabes. Par là les langues muṇḍā se distinguent à première vue des autres langues austroasiatiques et ressemblent aux langues dites agglutinantes telles que le turc.

A cet égard, les langues muṇḍā sont même plus avancées que la plupart des langues dites « agglutinantes ». Tout substantif ou verbe peut en effet s'adjoindre un ou plusieurs pronoms qui viennent se placer soit à la suite du mot, soit entre les morphèmes et la racine. Il en résulte des agrégats extrêmement complexes qu'on pourrait appeler « groupes pronominaux » et qui sont formés de la juxtaposition d'une racine principale, redoublée ou non, de divers morphèmes et de pronoms. Par exemple, en santālī la phrase *hâpân-iñ-e dal-ket'-ta-ko-tiñ-a* « mon fils a frappé le leur » est formée de deux groupes dont le premier *hapan-iñ-e* signifie littéralement « fils-de moi-lui » et le second *dal-ket'-ta-ko-tiñ-a* est composé du verbe *dal-ket'-a* « a frappé » et des pronoms *ta-ko* « d'eux » et *tiñ* « de moi » insérés entre les éléments du verbe. La phrase entière signifie donc littéralement « lui mon fils a frappé le leur, lui qui est à moi ». Le pronom *e* « lui » du premier groupe est emphatique ; le pronom *tiñ* « de moi » du second groupe est destiné à rappeler le sujet. De tels groupes paraissent dus à la multiplication des pronoms jugés nécessaires pour donner à la phrase plus de cohésion, de clarté et de relief comme il arrive fréquemment dans le langage populaire.

J. Przyluski.

LANGUES MALAYO-POLYNÉSIENNES

PRÉAMBULE

Les langues malayo-polynésiennes offrent les caractéristiques suivantes : le thème radical est généralement un dissyllabe paroxyton ; le nom n'a ni flexion, ni genre, ni nombre ; la catégorie du verbe s'exprime à l'aide d'une grande variété de morphèmes (préfixes, infixes et suffixes) qui forment des verbes transitifs ou intransitifs, causatifs, passifs, réciproques, potentiels, relatifs, fréquentatifs, etc. Les langues du domaine sont différenciées entre elles par des correspondances phonétiques constantes ; la morphologie et la syntaxe sont, dans l'ensemble, homogènes.

Les langues et dialectes des îles situées à l'Est du 120[e] degré de longitude de Greenwich, tout en conservant les caractères généraux de l'indonésien, montrent une influence mélanésienne ou papoue, tant dans le vocabulaire que dans la morphologie et la syntaxe, influence qui s'accentue au fur et à mesure qu'on approche de la Nouvelle-Guinée et des îles de la Mélanésie.

Le domaine des langues malayo-polynésiennes s'étend de façon discontinue, de Madagascar (40[e] degré de longitude orientale) à l'île de Pâques (110[e] degré de longitude occidentale), c'est-à-dire sur une étendue de 210 degrés de longitude (voir planche 12). Ce domaine exclusivement insulaire a été divisé, pour la commodité de l'exposition, en indonésien, mélanésien, micronésien et polynésien. Ces divisions ne correspondent qu'approximativement à la distribution des langues envisagées en familles et groupes ; mais cette répartition arbitraire, quelque désuète qu'elle soit à certains égards, peut être maintenue sans inconvénient ici, où il n'est pas possible de proposer et de justifier une classification nouvelle, strictement linguistique.

INDONÉSIEN

Généralités.

Créée par H. Neubronner van der Tuuk, la linguistique indonésienne dut son développement et sa méthode scientifique à Hendrik Kern qui y consacra sa vie. De nombreux élèves du maître, devenus maîtres à leur tour, ont étudié en Indonésie même un grand nombre de langues et dialectes, et leurs travaux font honneur à l'école de Leyde. En 1893, le Dr Renward Brandstetter de Lucerne commençait la publication de ses *Malaio-polynesische Forschungen* qui se poursuit sans interruption depuis cette époque. Ce sont ces travaux qu'on a pris pour guide et dont on a résumé ici les principales conclusions.

D'après l'*Encyclopaedie van Nederlandsch-Indië*, le domaine géographique de l'indonésien compte une cinquantaine de millions d'habitants, dont plus des deux tiers sont sujets de la Hollande. Les îles de Java et de Madura sont habitées par 27.000.000 d'hommes dont 20.500.000 parlent le javanais, 6.500.000 le sunda ou langue de la partie occidentale de Java, et 3.000.000, résidant dans l'île de Madura et à la pointe orientale de Java, le madura. La langue de l'île de Bali est parlée par 3.000.000 d'hommes; le malais, par 3.000.000; la langue de Atšeh (ou Atchin, au nord de Sumatra), par quelque 500.000 hommes. Le reste, c'est-à-dire les langues batak, redžaṅ, lampoṅ, nias, dayak, les langues et dialectes du centre de l'île de Célèbes, etc., par 3.000.000 d'hommes; le bugi et le makassar du Sud de Célèbes, par 1.250.000. Aux Philippines, le tagal est la langue de 1.500.000 habitants; le bisaya, de 2.750.000; le bikol de Luçon, de 500.000. Enfin, à Madagascar, environ 3.000.000 d'habitants parlent le malgache.

L'Indonésie ou Insulinde est ainsi appelée pour rappeler sa colonisation par l'Inde qui remonte vraisemblablement au IV[e] ou III[e] siècle avant notre ère[1]. L'influence étrangère a laissé des traces dans les monuments anciens, l'épigraphie et les langues des îles occidentales du domaine, notamment à Sumatra, Java, Bali, Bornéo. Elle diminue en raison directe de l'éloignement des autres îles de Java-Sumatra qui furent les centres les plus importants de culture indienne et elle disparaît aux environs de la Nouvelle-Guinée.

Répartition géographique.

Le domaine indonésien peut se répartir géographiquement en huit groupes insulaires : 1° le groupe des Philippines; 2° le groupe de Célèbes ; 3° celui de Bornéo ; 4° le groupe de Java, Madura et Bali ; 5° celui des langues de Sumatra ; 6° le groupe indonésien de la partie méridionale de la péninsule de Malaka ; 7° les dialectes malgaches de Madagascar ; et, enfin, 8° les langues et dialectes des îles à l'Est du 120[e] degré de longitude. Les limites extrêmes du domaine sont : au Nord, Formose et les îles Batanes ; à l'Est, les îles en face de la Nouvelle-Guinée ; au Sud-Ouest, les îles de la côte occidentale de Sumatra (Simalur, Nias, Mentaway) ; à l'Ouest, Madagascar.

Les langues et dialectes connus de ce domaine sont les suivants :

ILE FORMOSE. — Favorlañ et autres dialectes formosans.

ARCHIPEL DES PHILIPPINES. — Les indications sur ce groupe sont empruntées à la *List of native tribes of the Philippines and of the languages spoken by them* par Ferdinand Blumentritt (dans *Smithsonian report for 1899*, p. 527-547, Washington, 1901,

1. Voir mon mémoire *Le K'ouen-louen et les anciennes navigations interocéaniques dans les mers du Sud*, dans *Journal Asiatique*, XI[e] série, t. XIII et XIV, 1919.

in-8°); voir également *A pronouncing gazetteer and geographical dictionary of the Philippine islands*, Washington, 1902, in-8° :

Abakas, peuple païen vivant dans les forêts de Caraballo Sur, à Luçon. Leur langue s'est complètement tagalisée.

Adañ, peuple vivant sur la montagne de ce nom dans la province de Ilocos Norte (Luçon).

Agutaino, langue de l'île de Agutaya dans l'archipel de Cuyo (province des Calamianes).

Apayao, peuple guerrier chasseur de têtes, habitant la partie N.-O. de la province de Cagayan (Luçon).

Aripa, peuple habitant près de Naksipiñ et Tubañ (Luçon).

Bagobo, peuple païen et guerrier habitant au pied du volcan Apo (Dávao, à Mindanao).

Batan, langue des îles Batanes, au nord de Luçon, la seule de l'archipel possédant le phonème *tš*, d'après T. H. Pardo.

Bikol ou Vikol, langue des habitants de la presqu'île de Camarines (Luçon) et de quelques îles voisines.

Bilan ou plutôt Buluan, langue des habitants des îles Sarañani, au large de la pointe Sud de Mindanao.

Bisaya (appelé officiellement Visaya), langue parlée dans le groupe insulaire au Sud de Luçon. Elle se divise en plusieurs dialectes dont les plus importants sont ceux des îles de Sebu (Cebu) et de Panay. Les dialectes bisaya sont parlés par plus de 2 millions d'indigènes.

Bontok ou Bontok-Igorrot, langue parlée dans la province de Bontok (Luçon) par un peuple de chasseurs de têtes.

Cagayan = Kagayan, nom de peuple habitant le bassin du Rio Grande de Cagayan à Luçon. Leur langue est l'ibanag. Ils n'ont rien de commun avec les Cagayans de Mindanao qui parlent un dialecte bisaya.

Calamian = Kalamian, dialecte bisaya.

Calaua = Kalawa, nom de peuple habitant les vallées du Rio Chico de Cagayan (Luçon). Leur langue est appelée également itavés, plutôt itawi (voir ci-dessous).

Cancanai = Kankanay, dialecte igorrot (voir ci-dessus Bontok-Igorrot) passé dans le N.-O. de la province de Benguet (Luçon).

Caragas = Karagas, anciens habitants de la côte orientale de Mindanao, appelés ainsi de leur ville principale Caraga = Karaga et qui parlaient, sinon une langue spéciale, tout au moins un dialecte bisaya.

Catalanganes = Katalañanes, nom de peuple vivant dans le bassin de la rivière Catalangan (province de Isabela, à Luçon). Ils parlent la même langue que les Irayas (voir ci-dessous).

Cataoan = Kataoan, dialecte parlé par les Igorrots (voir ci-dessus Bontok-Igorrot) du district de Lepanto (vallée de la rivière Abra à Luçon).

Cebuan = Sebuan. Voir ci-dessus Bisaya.

Coyuvo = Koyuvo, dialecte de l'archipel de Cuyo = Kuyo (groupe insulaire de la mer de Mindoro) qui est sans doute un dialecte tagbanua. Les Tagbanuas sont une peuplade d'origine malaise métissée par des Negritos.

Dadayag, langue de sauvages païens et chasseurs de têtes des montagnes occidentales de Cagayan = Kagayan (province de Cagayan, dans la mer de Sulu).

Gaddañ, langue de chasseurs de têtes habitant les provinces d'Isabela et Cagayan, à Luçon.

Gamuñan ou Gamunañan, langue d'un peuple habitant la partie montagneuse à l'Est et au Nord de Tuao (province de Cagayan, à Luçon).

Giaña ou Guaña, langue d'un peuple habitant le Nord-Est et le Nord de Dávao (Mindanao).

Ginaan, langue de chasseurs de têtes habitant le bassin du Rio Abra et du Rio Grande (Luçon).

Halaya, dialecte bisaya parlé dans l'intérieur de l'île de Panay.

Haraya. Voir le précédent.

Hiligayna, dialecte bisaya parlé sur la côte de l'île de Panay.

Ibanag, langue parlée par les Cagayans. Voir ci-dessus sous Cagayan.

Ifugao, langue d'un peuple de chasseurs de têtes habitant les provinces de Nueva Vizcaya et Isabel (Luçon).

Igorrot, anciennement noté Igolot, langue d'un peuple païen

et guerrier habitant actuellement la province de Benguet (Luçon). Voir ci-dessus Bontok.

Ilocan = Ilokan, langue d'un peuple habitant les provinces de Ilocos Norte, Ilocos Sur, Union, à Luçon.

Inibaloi, dialecte parlé par les Igorrots Agnothales. Voir ci-dessus Igorrot.

Iraya, langue des habitants du bassin de la rivière Catalañan = Katalañan, à Luçon.

Isinai, langue parlée dans la Nueva Vizcaya (Luçon).

Itawi, dialecte gaddañ (voir ci-dessus Calaua).

Ivaña, dialecte ibanag. Voir ci-dessus Ibanag.

Magindanao, langue de Mindanao.

Mayoyao, dialecte de l'Ifugao (voir ci-dessus).

Pampaño, langue parlée dans la province de Pampaña, à Porac, et dans quelques îlots linguistiques des provinces de Nueva Ecija, Bataán, Zambales (Luçon).

Quiangan = Kiañan, dialecte ifugao (voir ci-dessus) parlé dans la commanderie de Quiangan (Luçon).

Silipan, dialecte ifugao (voir ci-dessus) parlé dans la province de Nueva Vizcaya et la commanderie de Quiangan (Luçon).

Tagal, langue parlée dans les provinces de Manille, Corregidor, Cavite, Bataán, Bulacán, Batangas, Infanta, Laguna (Luçon) et l'île de Mindoro; et, avec une moindre expansion, dans les provinces de Tayabas, Zambales, Nueva Ecija, Isabela et Principe, à Luçon.

Tiñyan, langue des indigènes habitant le Nord de Babayan Creek, dans l'île de Palawan.

Tino, langue de la province de Zambales (Luçon).

Tirurai, langue des montagnards de l'Ouest du Rio Grande (île de Mindanao).

Iles Sañi ou Sañir. — Sañir avec trois dialectes : Mañanitu, Tugulandañ et Siauw.

Ile de Célèbes. — Le groupe linguistique de l'île Célèbes comprend, du Nord au Sud, les langues et dialectes suivants :

langues ou dialectes de Tonséa dans les districts de Tonséa et Maumbi; Tombulu dans les districts de Tomohon-Saroṅsoṅ, Tombariri, Kakaskasen et Menado; Tombantik dans le district de Bantik; Tulur dans les districts de Tondano-Tulian, Tondano-Tulimambot et Kakas-Remboken; Tulaṅkuan dans le district de Laṅowan; Tumpakĕwa, actuellement appelé Tontemboan, dans les districts de Sonder, Kawaṅkoan, Tompaso, Tombasian et Romoön, avec deux sous-dialectes : Matanai et Makalai; Tunsini dans le district de Tonsawaṅ; Tumpasaṅbaṅko ou Tumbenténau et Tumbélaṅ, ces deux derniers dans le district de Pasan-Ratahan-Ponosakan ; Bolaaṅ-Moṅondow, Gorontalo.

Centre-Célèbes, sur le détroit de Makassar : Tawaili avec deux dialectes, Palu avec deux dialectes, langue de To Loli et de To Ganti; Sigi, Pakuli avec deux dialectes; à l'intérieur : Kulawi, Pobatua, Lindu; plus à l'Est : Parigi ou Tara, Sausu. Ces langues et dialectes sont désignés sous le nom de groupe Toradža occidental. Le groupe Toradža oriental comprend les langues suivantes : Napu, Besoa et Bada, auxquels s'ajoute le Bare'e avec les sous-dialectes Ampana, la langue des îles Togian, dans la baie de Tomini, To Lalaeo, To Rapaṅkaka; au Sud-Est du lac de Posso, Pu'u na Boto, To Salu-maoge, To Lampu.

Sud-Célèbes : Bugi et Makassar qui sont étroitement apparentés.

Bornéo. — Dayak et ses dialectes. Sur la côte orientale de l'île, M. Mervyn W. H. Beech a étudié sommairement le groupe Tidoṅ (*The Tidong dialects of Borneo*, Oxford, 1908) et en reconstitue ainsi la généalogie : Tidoṅ ancien représenté par deux dialectes vivants : Boloṅan et Tarakan; ce dernier se divise en plusieurs sous-dialectes : Simbakoṅ, Nonoekan, etc. Un dialecte spécial dit *Basa Saṅyaṅ* ou langue des esprits.

Java. — A l'Ouest, le Sunda; à l'Est, le Javanais; à l'extrême Est et dans l'île de Madura, le Madura. On distingue en javanais deux sortes de langues : le *kromo* ou langue officielle et le *ṅoko* ou langue vulgaire.

Il en est de même en sunda, à Madura et à Bali (voir page 412).

Sumatra. — Du Nord au Sud : Atšeh, vulgairement appelé Atšin ; Gayō ; Batak avec les dialectes karo-batak, toba-batak, dairi-batak, mandailiṅ-batak, dialecte batak d'Aṅkola, timur-batak, dialectes de Pane et une langue spéciale dite langue de anduṅ « complainte pour un mort » qui est une sorte de *vocero* ; Minaṅkabaw ; Redźan, Lampoṅ et Malais sur la côte orientale. La célèbre inscription de Kota Kapur, sur l'île de Baṅka, nous a conservé une proclamation officielle du VIIe siècle qui est rédigée en une langue à cérébrales représentant peut-être une sorte de vieux malais. C'était, ou ce devait être, la langue officielle en usage dans l'ancien empire de Çrīvijaya, le Che-li-fo-che des textes chinois, le Sribuza des Arabes, dont la capitale était à Palembaṅ à l'époque où fut rédigée l'inscription de Kota Kapur.

Sud de la presqu'île de Malaka. — Malais du continent et des petites îles voisines (Riouw, Liṅga, etc.). Au XVIe siècle, dans les recueils chinois de vocabulaires étrangers appelés *houa-yi yi-yu*, le malais est désigné sous le nom de *Man-la-kia* = Malaka (voir *Bulletin École française d'Extrême-Orient*, t. XII, 1912, fasc. 9, p. 198 et suiv.).

Iles de la côte occidentale de Sumatra. — Simalur ou Simölu, Nias, Mentaway.

Madagascar. — Les principaux dialectes de la grande île africaine sont :

Côte orientale, du Nord au Sud : Antankára, Betsimisáraka, île de Sainte-Marie, Ranuména, Antambahwáka, Antaimúru, Antaifáši, Antaišáka, Antaimanambúndru, Antanóši, Antandrwí.

Côte occidentale, du Sud au Nord : Mahafáli ou Mafáli, Vézu, Fieréṅa, Menabé, Sakaláva.

Intérieur de l'île, du Nord au Sud : Sakalava Nord-Est, Antsihánaka, Mérina (improprement appelé Hova), Bezanuzánu, Vurímu, Betsileo, Antaṅála, Bára.

Iles a l'est de Java. — Ce domaine se divise ainsi :

Bali. La langue de Bali se divise, comme le javanais, en haut

et bas balinais ou langue officielle et langue vulgaire. Celle-là est très voisine du javanais *kromo*; celle-ci est plus proche du malais et de la langue de Sunda (javanais occidental) que du javanais propre.

LOMBOK, SUMBAWA.

Les îles suivantes constituent un groupe linguistique distinct qui comprend :

L'île de FLORES ou ENDE (exactement *Hende*), l'île de SUMBA avec deux dialectes : Sumba oriental et Laora ou Sumba occidental; SAWU; l'île de ROTTI avec les dialectes de Termanu ou Pada, de Baä, Bilba, Bokai, Korbaffo, Deṅka, Riṅgow ou Reṅgow, Ti et U-nale.

TIMOR avec les dialectes principaux Timol, Belu, Helu, et les sous-dialectes Tapenu ou Taibeno, Sonabai-ana, Oëmatan ou Amakono et Dženilo. Dans la partie orientale de l'île qui appartient au Portugal, on parle le dialecte Teto.

LETTI ou LETI, MOA, LEIKOR ou LAKOR, ROMA et KISAR ou KISSER, appelés groupe de Letti.

BABAR et îlots voisins.

TANIMBAR ou TENIMBAR appelé également TIMORLAUT et îlots voisins.

ARU et îles voisines.

Groupe insulaire de BANDA.

Groupe d'AMBOINE : parlers des 4 *negori* ou villes (Seit, Lima, Uriṅ, Asilulu); parlers de Hila (Hila, Wakal, Hitulama, Hitumesiṅ, Mamala, Morela); parlers de Alaṅ, Hatu, Liliboi, Batumerah, Kaitetu, Wakasihu, Liaṅ et Tulehu.

HARUKU, NUSA LAUT et SAPARUA.

Ile de CERAN ou SERAN. La partie orientale de cette île est habitée par une population d'origine indonésienne occidentale, d'une part; papoue, de l'autre. Un groupe linguistique de la côte méridionale de l'île (baie de Elpupati) a été récemment bien étudié par M. Erwin Stresemann (*Die Paulohisprache. Ein Beitrag zur Kenntnis der Amboinischen Sprachengruppe*, Leyde, 1918, in-8°). Le système consonantique du Paulohi ne comprend que 11 consonnes : *p*, *f*, *v*, *h*, *l*, *ꝛ*, *s*, *t*, *k*, *m*, *n*. Les pronoms personnels

distinguent, comme en indonésien, l'inclusif et l'exclusif à la 1re pers. du sing. et du pluriel; il y a une double forme pronominale : la première pour les personnes et la seconde pour les animaux et les choses. L'ouvrage précité contient une grammaire et un vocabulaire comparés du paulohi avec les langues et dialectes de l'indonésien, du mélanésien et des langues de la Nouvelle-Guinée.

Archipel des MOLUQUES. Les langues et dialectes de l'île de Halmahera et des îles voisines (Ternate, Tidore, Motir, Makiyan ou Makian) ne sont pas suffisamment connus pour être classé avec certitude dans l'indonésien.

Histoire littéraire.

En fait, le javanais est l'unique langue de l'indonésien dont nous puissions reconstituer l'histoire à relativement haute époque. La langue ancienne ou vieux-javanais est également appelée *kavi*, litt. « la langue du poète » ; les plus anciens témoignages remontent aux environs de l'an 800 de notre ère. Le vieux-javanais se différencie du javanais moderne surtout par sa phonétique et sa morphologie. Vieux-javanais *abwat* « lourd », par exemple, est passé dans la langue moderne à *abot* par contraction de *w* + *a* en *o*.

Phonétique.

Le système phonétique de l'indonésien commun se composait vraisemblablement de :

6 voyelles : *a*, *i*, *u*, *e*, *o*, *ĕ*.
2 semi-voyelles : *y*, *w*;
19 consonnes : 3 gutturales : *k*, *g*, *ṅ*;
3 prépalatales : *tš*, *dž*, *ñ*;
3 dentales : *t*, *d*, *n*;
3 labiales : *p*, *b*, *m*;
3 liquides : *r*, *ϱ*, *l*;
1 sifflante : *s*;

1 aspirée : *h*;
1 occlusive glottale : '.

La voyelle transcrite conventionnellement par *ĕ* est une voyelle indéterminée qui a une certaine analogie avec *ö* allemand; elle est généralement désignée sous son nom javanais de *pĕpĕt*, à peu près **pöpöt*. Cette voyelle n'est représentée que dans un relativement petit nombre de langues du domaine : bugi, *tĕñña*; bolo-ñan, sunda, karo-batak, malais, simalur, *tĕñah*; sawu, *tĕña*. Les autres langues et dialectes y répondent par l'une des cinq autres voyelles : indonésien commun *tĕlú* « trois » > pañasinan *taló*, malgache *télu*, sunda *tílu*, toba-batak *tólu*, tingi *túlu*. Un traitement identique du *pĕpĕt* s'observe dans les transcriptions de noms indonésiens : *ĕ* a été rendu indifféremment par *a*, *e*, *i* ou *u*. Ces divergences peuvent s'expliquer par l'existence dans les langues modernes d'une articulation différente du *pĕpĕt* tendant vers *a*, *i* ou *u*. Cette dernière articulation semble attestée par l'équation : indonésien commun *bĕẓat* > vieux-javanais *bwat*, alors qu'on attendait **bĕat* par chute régulière de *ẓ* commun dans cette dernière langue.

Dans quelques langues, le *pĕpĕt*, comme les autres voyelles, peut être long ou bref, tonique ou atone. En dehors de ces cas exceptionnels, il est, au contraire, bref et atone. Voir par exemple l'opposition en gayo entre *túluk* « vérifier » et *tĕlúk* « baie ».

Aux cinq autres voyelles de l'indonésien commun se sont ajoutées les voyelles nasales et, dans quelques langues et dialectes dont le dayak, le makassar, le bontok des Philippines, etc., les trois voyelles *ä*, *ö* et *ü*, et, en vieux-javanais, *ṛ*, *ṝ*, *ḷ*, *ḹ* qui ont été empruntés aux langues de l'Inde.

Le système consonantique de l'indonésien commun s'est montré d'une grande stabilité dans un aussi vaste domaine insulaire. On constate, sans doute, d'importantes variations : ainsi, *d* malais est une sorte de cérébrale que le javanais, qui possède un *d* dental et un *ḍ* cérébral, rend toujours par *ḍ*; *tś* est alvéolaire en javanais, palatal en malais ; *w* est dentilabial en bunku ; *f* bilabial en buli de Halmahera des Moluques ; mais ce sont des divergences phonétiques attendues, et on s'étonne presque qu'elles ne soient

pas plus nombreuses. D'autre part, si quelques phonèmes ont sporadiquement disparu (ꝑ en vieux-javanais; *r* et les palatales en rotti), de nouveaux apparaissent, soit par emprunt à des langues étrangères, soit par évolution naturelle d'un phonème indonésien. Le kavi ou vieux-javanais a, par exemple, emprunté à l'Inde ses gutturales, palatales, dentales et labiales aspirées ; ses cérébrales pures et aspirées ; les sifflantes *ç* et *ṣ*, l'*anusvāra*, le *visarga* et l'*anunāsika*.

On relève, en outre, la vélaire *ḫ* en nias, la spirante *g* en sañir ; la sifflante palatale *ś* à Madagascar, Mentaway et aux Philippines.

Les dialectes de Madagascar fournissent les nouveaux phonèmes suivants :

ẑ < malgache ancien *dẑ* < indonésien occidental *dẑ*;

dz[1] < indonésien commun **dẑ*;

dr et *tr* qui s'articulent à peu près comme les *dr* et *tr* anglais de *drive, tree* (telle est leur prononciation en merina ; les mêmes phonèmes dans les autres dialectes, et particulièrement dans les dialectes sud-orientaux, se prononcent un peu plus en arrière qu'en merina) ;

l'affriquée *ts* (l'affriquée *tš* était représentée en malgache sud-oriental ancien ; je n'en connais qu'un seul exemple dans les dialectes orientaux modernes);

la spirante *z* issue de *yod* ou de *dẑ* indonésien commun ;

les dentilabiales *f* et *v* issues de *p* et *w* ou *b* de l'indonésien commun (*f* est également attesté à Rotti et aux Philippines);

enfin des phonèmes avec nasale antérieure, tels que tontemboan *ṅk*, *mp*, *mb*, *nt*, *nd*, *ntš* où la nasale fait partie intégrante de la consonne suivante.

LOI ꝑ, *g*, *h*. — A l'ꝑ uvulaire de l'indonésien commun, le bugi, le makassar, le malais, le toba-batak répondent par ꝑ ; le tagal et le bisaya par *g*, le dayak par *h*, le vieux-javanais par zéro ; d'autres langues par *r*, *l*, *y*, *z*, ', *k* > *tš*, *h*, *t*; *g* : ind. com. **uꝑat*

1. En graphie usuelle *j* ; de même que *o* de l'orthographe courante = *u* et *y* final = *i*. Pour ne pas dérouter le lecteur, on a maintenu l'orthographe officielle sur la carte qui accompagne ce texte (planche 12).

« veine » > malais *uꝫat*, tagal *ugát*, dayak *uhat*, pañasinan *ulát*, lampoṅ *uyat*, malgache *úzatra*, tontemboan *o'at*; ind. com. **niuꝫ* « noix de coco » > talaut *níuka*, variantes d'autres parlers talaut : *niutša*, *níuha*, *niuta*; ind. com. **bĕꝫas* « riz » > saṅir *bóɣasĕ'*; ind. com. **dža꞉um* « aiguille » > madura *džharum*.

Liquides. — R lingual s'est maintenu dans presque tout le domaine : indonésien commun **pira* « combien ? » > vieux-javanais *pira*, malgache *píri* (comme second terme d'un complexe), *fíri*; mais dans un certain nombre de parlers il est modifié : il passe à *l* : bisaya *pila* « combien ? », à *d* : bali *pidan* « combien ? »; quelquefois à *g* et à *h* : ind. com. *iruṅ* « nez » > toba-batak *iguṅ*, nias *ihu*; à zéro : ind. com. **butir* « grain, globule » > malgache ancien *buti*, malg. mod. *vutsi* « verrue ».

L s'est généralement maintenu dans le domaine : indonésien commun **laṅit* « ciel » > vieux-javanais, malais, sunda, madura batak, dayak, bisaya, etc., *laṅit*; malgache oriental *laṅitsə*, *laṅitsi*, *laṅitrə*, merina *lanitra*. Mais il peut aussi être altéré : il passe à *r* : ind. com. **lapar* « faim » > toba-batak *rapar* par assimilation à la finale; à *y* à l'intervocalique : ind. com. **džalan* « chemin » > bare'e *džaya*; à *w* en tagal : ind. com. **puluh* « dix » > *powo*; à *n* en timor : ind. com. **kali* « graver » > *hani*; à *d* : ind. com. **lima* « cinq » > merina *dimi*; à *g* : ind. com. **ulu* « tête » > batan *ogo*; à *h* dans les dialectes de Formose : *uho* « tête ».

Lois des prépalatales. — Indonésien commun **tš* se maintient dans certaines langues : ind. com. **ratšun* « poison » > vieux-javanais, malais, etc., *ratšun*, bima *ratšu*; passe à *s* dans d'autres : tagal *lason* « poison ».

Indonésien commun **dž* s'est maintenu en nombre de cas : *džalan* « chemin » > bontok, bĕsĕmah *džalan*, bare'e *džaya*. Il est passé à *tš* en bugi, après *ñ* : ind. com. **džañdži* « promesse » > bugi *džañtši*; à *dž* aspiré en madura : ind. com. **džalan* > mad. *džhalan*; à *d* en vieux-javanais : *dalan* « chemin »; à *l* et *z* en malgache : *lálana* « chemin »; ind. com. **tudžu* > malg. *tūzu*; à *s* en lalaki : ind. comm. **džalan* > *sala* « chemin »:

Indonésien commun **ñ* s'est maintenu ou est passé à la nasale

dentale : ***pĕñu* « tortue » > vieux-javanais, malais, etc. *pĕñu*, madura *pĕñño*, batak *ponu*, malgache *fañu*, *fanu*.

Lois des dentales. — *T* de l'indonésien commun se maintient dans la plupart des langues et dialectes du domaine : ind. com. **tali* « corde » > *tali* en vieux-javanais et javanais moderne, malais, batak, dayak, makassar, bugi, malgache sud-oriental (*tadi*, dans les autres dialectes), etc. ; passe quelquefois à la sonore : ind. com. **mata* « œil » > sawu *mada* ; à *ts* en malgache devant *i* : ind. com. **tilik* « regarder » > dialectes malgaches *tsidika*, *tsilika* ; à *ḫ* : ind. comm. **tai* « excrément » > dialectes de Formose *ḫee* ; à *h* : ind. comm. **pitu* « sept » > kamberi *pihu* ; à *s* après *i* en bolaań-mońondow : ind. com. **kulit* « peau » > *kulis* ; à l'occlusive glottale en bugi : *uliʾ* ; à zéro : ind. com. *kulit* > nias *uli*.

D de l'indonésien commun se maintient en beaucoup de langues : ind. com. **dagań* « étranger » > *dagań* en vieux-javanais, toba-batak, etc. ; *daga* en bima ; passe à la sourde à la finale : ind. com. **añud* « courant (Strömung) » > malais *hañut* ; à *dh* ou *d* aspiré : ind. com. **damar* « résine » > madura *dhamar* ; à *r* après *n* : ind. com. **linduń* « ombre » > bugi *linruń* ; à *dr* après *n* : ind. com. **tanduk* « corne » > malgache *tandruka* ; à zéro : kulawi *tonu* « corne ».

En plusieurs langues, dont le madura, *t* et *d* de l'indonésien commun sont quelquefois cérébralisés.

N de l'indonésien commun se maintient dans le plus grand nombre des langues du domaine ; passe fréquemment à *ń* à la finale : ind. com. **anin* « vent » > bugi *añiń* ; à *l* : ind. com. **anak* « enfant » > dialectes de Formose *alok* ; s'assimile dans la langue parlée à un *t* subséquent : ind. com. *gantuń* « pendre, être suspendu » > toba-batak parlé *gattuń*, mais la langue écrite reproduit le thème de l'indonésien commun ; passe à zéro devant *t* : ind. com. *lintah* « sangsue » > nias *lita*.

Lois des labiales. — *P* se maintient dans la majorité des langues du domaine ; passe fréquemment à la dentilabiale *f*, quelquefois à *h* : ind. com. **pitu* « sept » > vieux-javanais *pitu*, malgache *fitu*, rotti *hitu*. Dans quelques tribus Toba-Batak de l'Est, la labiale sourde des autres dialectes batak, étant devenue

imprononçable, a noté Van der Tuuk, est passée à *k* : toba-batak *piso* « couteau » > toba-batak oriental *kiso*. Enfin, *p* passe à zéro : ind. comm. **pira* « combien ? » > kissar *ira*.

B de l'indonésien commun est généralement maintenu. Il est aspiré en certains cas : ind. com. **buru* « chasser (jagen) » > madura *bhuru* ; passe à *p* : ind. com. *bulu* « cheveu » > buli *plu* ; à *v* : ind. com. **bañaw* « héron » > malgache *vanu* ; à *f* : ind. comm. **ribu* « mille » > rotti *lifu* ; à *h* : ind. com. **bĕli* « prix » > makassar *balli* > dialecte silayar du makassar *halli* ; à zéro : ind. com. **batu* « pierre » > gayo *atu*.

L'ASPIRÉE *h* de l'indonésien commun se maintient dans quelques langues : ind. com. **pĕnuh* « plein » > vieux-javanais *pĕnuh*, tarakan *panuh* : Elle passe à l'occlusive glottale en certains cas : ind. com. **lintah* « sangsue » > tontemboan *linta'* ; à zéro : bugi *pĕnno* « plein ».

GÉMINATION DES CONSONNES. — Elle ne se présente que sporadiquement, dans un nombre restreint de langues de la famille et dans des conditions déterminées. Ainsi, par exemple, en bugi et en makassar après le pĕpĕt ou la voyelle qui répond au pĕpĕt de l'indonésien commun : ind. com. **tĕkĕn* « bâton » > vieux-javanais et javanais moderne *tĕkĕn*, mais bugi *tĕkkĕn*, makassar *takkaṅ*.

LE SANDHI. — Ce mot sanskrit qui signifie au propre « articulation, jointure » et, en grammaire, « juxtaposition euphonique des mots », est passé en indonésien sous les formes *sĕndi*, *sandi*, avec ce dernier sens. Dans un complexe, la loi de sandhi s'exerce tant sur la consonne finale du premier mot que sur la consonne initiale du mot suivant : malais *paṅgal* « coupé » ; *mĕmaṅgal* « couper » < *mĕṅ* + *paṅgal*, malgache *tafika* « expédition de guerre », *manafika* « partir en expédition « < *man* + *tafika*, *ulikazu* « vers du bois » < *ulitra* + *hazu*.

Morphologie.

LE THÈME RADICAL indonésien est quelquefois monosyllabe ou dissyllabe oxyton, mais plus généralement dissyllabe paroxyton. Par exemple : vieux-javanais *uṅkab*, *liput* ; karo-batak *liṅgĕm*,

tutuñ, *lintañ*, *idah*. Une enquête à l'intérieur du vieux-javanais et du dialecte batak permet de se rendre compte qu'il s'agit ici non pas d'un véritable thème radical, mais d'un thème radical secondaire :

vieux-javanais : *uñkab* « ouvrir », *siñkab* « découvrir » < *kab* ; *liput* « couvrir », *saput* « couvrir » < *put* ;

karo-batak : *liñgĕm* « ombre », *agĕm* « ciel nuageux » < *gĕm* ; *lintañ* « bandes, stries », *rintañ* « rangée », *listañ*, *liñtañ* « en ligne droite » < *tañ* ;

idah « voir », *dedah* « regarder », *tšidah* « montrer », *pĕdah* « apprendre, enseigner » < *dah*.

La comparaison entre les différentes langues du domaine fournit des indications identiques :

tagal, bisaya *hasañ* ; bulu, tontemboan *asañ* ; gayo *isañ* ; malais *insañ* « branchies » < *sañ* ;

bali *sĕlsĕl* ; vieux-javanais *sĕsĕl* ; tonsea *manĕsĕl* ; bisaya *basol* ; malais *sĕsal* ; makassar *sassala'* « repentir » < indonésien commun **sĕl*.

Ces exemples qu'on pourrait multiplier, montrent que si l'indonésien commun que nous pouvons restituer était généralement dissyllabique, ce dissyllabisme révèle à l'analyse un monosyllabisme antérieur.

Les emprunts. — Les langues indonésiennes ont fait de nombreux emprunts au sanskrit, à l'arabe, au portugais et au hollandais ; et de moins importants au persan, aux langues dravidiennes de l'Inde, au chinois, etc... Le sanskrit a fourni des mots de civilisation et des noms de produits étrangers ; l'arabe, des termes de religion et de droit mulsumans; le portugais, des termes usités dans le commerce et l'industrie ; le hollandais, des expressions techniques administratives, maritimes et militaires. Les emprunts du kavi ou vieux-javanais au sanskrit sont en nombre si considérable que le *Rāmāyaṇa* kavi contient des phrases de ce type (IV, 9) : *tĕmu ñ-āçrama dibya çobha rāmya* « [il] atteignit un bel ermitage, propre et plaisant », où le premier mot et l'article *ñ* sont seuls indonésiens. Mais M. Brandstetter qui

a cité cet exemple caractéristique fait justement remarquer que la littérature kavi est une production artificielle reproduisant servilement les textes de l'Inde, au lieu de les traduire au sens exact du mot. Il est évident que le vieux-javanais parlé devait être moins hindouisant que la langue littéraire et que son caractère indonésien devait être beaucoup plus accentué.

LE VERBE. — Dans la phrase suivante qui est empruntée à l'hymne des héros de l'île de Nias : *moi mura'u dotoa hulayo* « [il] alla [et] saisit [le] bois [de la] lance », *moi* est le verbe « aller » ; ce mot n'a pas et ne peut avoir d'autre sens, il exprime l'idée de « se mouvoir, aller ». *Mura'u*, au contraire, ne signifie « saisir, embrasser » que grâce au morphème verbal *mu* préfixé au thème radical *ra'u*. De même dans ce passage de l'épopée malaise *Bidasari* ; *tiyap tiyap hari dudu' bĕrtšinta* « chaque jour, [il] s'asseyait [là] attristé », *dudu'* est le verbe « s'asseoir » et ne peut signifier autre chose. Ainsi l'indonésien possède deux sortes de verbes : un verbe simple, autrement dit un thème radical à sens verbal du type *moi*, *dudu'* ; et un verbe composé où le sens verbal du complexe est indiqué par un morphème : *mura'u* < préfixe verbal *mu* + thème radical *ra'u*.

Le verbe simple est pour l'indonésien un verbe et rien de plus. A l'analyse des textes par un linguiste européen, apparaissent les différences de sens que voici : dayak : *blaku tiñak* « [il] désire être parlé, il désire qu'on parle » ; kamber : *na hadañ* « il se lève » ; vieux-javanais : *wĕruh ri ambĕk* « être familier avec la vie intérieure » ; malais : *tahu bahasa* « posséder [de] l'instruction ». Des quatre verbes simples précédents, le premier, *tiñak*, a un sens passif ; le second, *hadañ* est un neutre ; le troisième, *wĕruh*, qui signifie « être familier (avec quelqu'un ou quelque chose), est intransitif et doit user d'une préposition — dans le cas présent *ri* — pour se relier au complément ; le quatrième, *tahu*, a un sens actif. Les verbes simples de la première et de la seconde sorte sont les plus nombreux ; ceux de la troisième sorte sont plus rares et ceux de la quatrième extrêmement rares.

D'après les exemples précédents, on voit qu'aucune marque

extérieure n'indique à priori le caractère soit verbal soit nominal d'un thème radical donné. Malais *dudu'* signifie « s'asseoir », et non « siège » ; en bugi, *api* est « feu » et *tunu* « brûler », les deux mots étant étymologiquement étrangers l'un à l'autre. D'autre part, dans quelques cas, le même mot réunit les deux sens, nominal et verbal. C'est, par exemple, le cas de malais *tidor* qui signifie également « dormir » et « sommeil » : *tidor iya* « il dort », *tidor ña* « son sommeil ». Celui-là sert de verbe ainsi que l'indique le pronom personnel sujet *iya*; celui-ci, de substantif, avec possessif *ña* suffixé.

L'indonésien connaît trois sortes de verbes composés : actif, causatif et passif.

Les morphèmes de l'indonésien commun pour le verbe actif sont : *ma-*, *ṅ-* ou *maṅ-* < *ma* + *ṅ*, et *um-* préfixe, quand l'initiale du thème radical est une voyelle ; — *-um-* infixe, quand l'initiale est une consonne.

Morphème actif *ma* : Formose *ma-tagga* « saigner », bontok *ma-suyep* « dormir », bunku de Célèbes *mahaki* « être malade », *'Basa Saṅiaṅ* de Bornéo *ma-hampan* « traverser », bali *ma-humah* « habiter », kamber *ma-lala* « cuire », lampoṅ *ma-barsog* « nasiller », mentaway *ma-loto* « avoir peur », malgache *ma-tahutra* « avoir peur ».

Morphèmes *ṅ* et *maṅ* : bontok *managni* (*maṅ* + *sagni*) « danser », bugi *ṅanro* (*ṅ* + *kanro*) « demander, solliciter », badžo *ṅinum* (*ṅ* + *inum*) « boire », *Basa Saṅiaṅ* de Bornéo *ṅ-udžan* « pleuvoir », javanais moderne *ṅ-utus* « envoyer », sumbawa *ṅ-adži* « enseigner », karo-batak *ṅ-apit* « serrer », mentaway *maṅarai* (*maṅ* + *karai*) « grimper, gravir », malgache ancien *nilu* < *ṅilu* « resplendir », malgache ancien et moderne *maṅ-araka*, merina *man-araka* « accompagner, suivre ».

Morphème *um* préfixe ou infixe : Formose *komma* (*ka* + infixe *um*) « parler », bontok *um-inum* « boire », tontemboan, dayak *k-um-an* « manger », vieux-javanais *k-um-ĕmit* « veiller », toba-batak *s-um-uruṅ* « être meilleur », simalur *l-um-aṅoi* « nager », malais *g-um-ilaṅ* « briller », malgache *h-um-ana* « manger ».

Ces exemples attestent le caractère indonésien commun des

trois morphèmes verbaux actifs. Dans un grand nombre de langues *ma-* forme des verbes intransitifs, *ṅ-* ou *maṅ-*, des verbes transitifs; dans d'autres langues, ces trois morphèmes forment un verbe actif dont le caractère transitif ou intransitif reste imprécis. Le morphème *um* en fonction d'infixe, exprime plutôt un sens aoristique, inchoatif ou futur : bontok *s-um-aa tša is nan fobfüy* « ils partent vers leur résidence », tontemboan *sia mĕla'ula'us omai t-um-aṅka' isera* « il arriva vite pour saisir ces choses-là », vieux-javanais *kagyat deni ṅ paksi madyus k-um-ĕlab* « effrayés par les oiseaux qui se baignaient, [les poissons multicolores] s'agitèrent ».

Aux trois morphèmes verbaux actifs *ma*, *ṅ* et *um*, s'en ajoutent d'autres qui sont d'un usage plus restreint et ont une moindre expansion.

Les deux principaux morphèmes passifs sont *ta-* et *in-* ou *-in-* (préfixe avec un thème radical à voyelle initiale, infixe avec thème à consonne initiale) : bisaya *ta-kilid* « être incliné », bunku de Célèbes *ta-peha* « être brisé », tarakan *ta-dagu* « être parlé », vieux-javanais *ta-wurag* « être dispersé », sawu *ta-bolo* « être plongé », toba-batak *ta-lentes* « être ouvert », mentaway *ta-itšo* « être vu », merina *ta-buruaka* « être transpercé »;

tagal *t-in-āwag* « être appelé », bulu de Célèbes *w-in-unu* « être tué », boloṅan de Bornéo *dẓ-in-awal* « être perdu », vieux-javanais *in-ambah* « être entré dans », kupaṅ *in-ka* « être mangé », lampoṅ *t-in-abor* « être répandu », mentaway *t-in-ibo* « être séché », malgache *t-in-apaka* « être brisé ».

D'autres morphèmes passifs sont également usités, tels que *ka-*, *taʔ-* (bugi *tā'*), le suffixe *-an*, etc. Le dayak possède cinq morphèmes passifs : *ba-kuntši* « être fermé », *i-agah* « être conduit », *ta-lenteṅ* « être coupé », *tar-adẓar* « être instruit », *tapa-isä* « être compté ». Au morphème passif du dayak, *tapa*, le malgache répond régulièrement par *tafa-* : *tafa-latsaka* « être tombé ».

Les modes. — L'impératif s'exprime de façons diverses. Tout d'abord, par des thèmes radicaux qui n'ont que cet unique sens : nias *aine* « viens ! », karo-batak *ota* « allons ! », dayak *hua* « fais attention ! » ; par d'autres thèmes qui répondent à latin *ecce*, tels que bontok *nay*, nias *hiẓa*, merina *indru*, etc.

L'impératif actif est, dans certains cas, identique au thème radical dont il n'est différencié que par l'intonation du sujet parlant (dayak *duan* « action d'aller chercher, va chercher ! »).

Il en est de même pour les verbes à morphème actif : mentaway *manuba* (*mañ-tuba*) « pêcher, pêche ! » Le morphème *pa-*, ordinairement causatif, exprime également l'impératif dans certaines langues : vieux-javanais *pa-mĕkuli* « embrasse ! », magindanao *pa-gedam* « lève-toi ! ».

Un grand nombre de langues indonésiennes possèdent une double négation dont l'une s'emploie exclusivement pour l'impératif : malgache *tsi mandzaka izi* « il ne gouverne pas » ; mais *aza mandzáka ami ni nufu, mandzaká ami ni fanahi* « ne gouverne pas avec la chair, [mais] gouverne avec l'esprit ». Le vétatif est exprimé par la particule négative *aza* < vieux-javanais *adža*, avec le présent *mandzáka*, et l'impératif positif sous sa forme habituelle *mandzaká* avec rejet de l'accent tonique à la syllabe posttonique du présent.

Quelques rares langues du domaine ont un morphème du subjonctif: vieux-javanais *-a*, bontok *-ed* après consonne, *-d* après voyelle ; tontemboan *-um-*. Les autres langues marquent ce mode à l'aide de particules, par exemple, malais *barañ* « il est possible que ».

L'optatif s'exprime soit par l'impératif, soit par des particules telles que *ma*, *mama*, *lah*, *malah*, etc.

Le morphème *maka-* forme des verbes potentiels dans un grand nombre de langues, notamment aux Philippines, à l'île Célèbes et à Madagascar (*maha-* < *maka-*).

Pour marquer qu'on veut que telle action s'accomplisse, on emploie soit le futur, soit, en gayo, par exemple, la particule *male* : *aku male ulak* « je veux [m'en] retourner ». *Devoir*, *pouvoir* se rendent par les expressions : *il faut*, *il est bon*, *il est convenable* : toba-batak *na so tupa mago hudžur Siriñis*, « litt. [il n'est] pas bon [que] devienne perdue [la] lance [de] Siriñis, il ne faut pas que la lance de Siriñis se perde, la lance de Siriñis ne doit pas se perdre ».

Les temps. — Voir ci-dessus, p. 421, pour le présent qui est exprimé par le thème radical auquel est préfixé le morphème actif. Le prétérit s'indique par un des morphèmes *ni-*, *no-*, *in-* ou *-in-* :

	Présent	Prétérit
Formose	*lummis* « brûler »	*l-in-ummis*
Bontok	*umdžanak* « arriver »	*in-umdžanak*
Talaut	*umire* « incliner la tête »	*in-umire*
Tontemboan	*maali* « apporter »	*ni-maali*
Nias	*mofano* « sortir »	*no-mofano*
Merina	*tunena* « être calme »	*nu-tunena*

En talaut, *umire* n'est pas un présent, mais un futur. Le morphème du prétérit *no-* à Nias, *nu-* en merina a, en toba-batak, les formes développées *nuṅ* et *nuṅa*. Le prétérit s'indique également par des particules indiquant le passé : bugi *pura*, makassar *lèba'*, vieux-javanais *huwus*, kupaṅ *hidi*, etc.

Le futur actif s'exprime diversement suivant les langues : par *h* < *hu* en malgache, par le préfixe *pi-* et le suffixe *-ön* à Sunda, par le morphème aoristique *um-* en tontemboan et bontok, par le présent, par le morphème du subjonctif *-a* en vieux-javanais, par des prépositions, par des sortes d'auxiliaires tels que malais *hĕnda'*, bontok *issa*.

Dans quelques langues, le trois temps du verbe et l'impératif constituent un système de conjugaison soumis à des règles invariables :

	Merina	Tagal
thème radical	*saruna* « couvercle »	*tawag* « appeler »
présent actif	*manáruna* « couvrir »	*tuṅmatawag*
prétérit	*nanaruna*	*tuṅmawag*
futur	*hanaruna*	*tatawag*
impératif	*manarúna*	*tumawag*

Temps du verbe causatif merina :

	Causatif
thème radical *tundra* « action de porter »	
présent *mitúndra* « porter »	*mampitúndra*

prétérit *nitundra* — *nampitundra*
futur *hitundra* — *hampitundra*
impératif *mitundrá* — *mampitundrá*

Temps du verbe passif en merina et tagal :

Merina : thème radical : *amé* « don, présent »
présent *uména* « être donné »
prétérit *numena*
futur *humena*
impératif *uméu*

Tagal : thème radical *tawăg* « appeler »
présent *tinatawag* « être crié »
prétérit *tinawag*
futur *tatawagin*
impératif *tawagin*

Le prétérit passif s'exprime en bugi par *pura*, en makassar par *le'ba'* suivis du thème radical. Ces deux sortes d'auxiliaires signifient littéralement « fini, terminé ». Voir par exemple, en bugi :

thème radical *siyo'* « lier »
prétérit actif *pura ma'siyo'*
prétérit passif *pura siyo'*.

Absence de flexion. — Le thème radical est invariable. Les personnes sont indiquées (quand il n'y a pas de substantif sujet) par antéposition ou postposition du pronom isolé, ou suffixation du pronom de la 1re, de la 2e ou de la 3e personne (soit pronom sujet, soit complément du verbe passif).

Le nom. — On vient de voir que l'indonésien connait deux sortes de verbes : le verbe simple ou thème radical à sens verbal, et le verbe composé, c'est-à-dire un thème radical à sens nominal ou verbal qu'un ou des morphèmes transforment en verbe actif, causatif ou passif. Parallèlement, le nom est simple ou composé. Le nom simple est un thème radical, généralement un dissyllabe paroxyton, dont rien n'indique la catégorie à l'état isolé, car le nom indonésien est invariable et n'a ni genre ni nombre. On a déjà montré, p. 422, que malais *tidor* a le double

sens verbal et nominal, et c'est la phrase seule qui déterminera l'un ou l'autre : employé avec un pronom personnel sujet, il faudra entendre *tidor iya* « il dort » ; avec le suffixe possessif : *tidor ña*, *tidor* est en fonction de nom et il faudra traduire « son sommeil ». Le cas de *tidor* reste heureusement une exception dans le domaine ; mais, même avec des thèmes différents, il y a parfois des difficultés d'interprétation : par exemple, malgache *hítani* et *tranuni* sont au premier abord embarrassants. Dans les deux cas, la finale *-ni* est le suffixe pronominal de la 3ᵉ personne du singulier qui est usité tant avec le nom qu'avec les verbes passifs. La morphologie du malgache ne fournit donc aucune indication précise pour distinguer lequel des deux mots est un nom et l'autre un verbe. Ce ne sera que le contexte qui permettra de traduire *hitani* par « il voit (litt. vu [*hita*] par lui) » et *tranuni* par « sa maison (litt. maison [*tranu*] de lui) ».

Le cas du nom composé est, au contraire, très clair : des morphèmes nominaux indiquent que les mots auxquels ils sont affixés appartiennent à la catégorie nominale : le préfixe *ka-* forme généralement des noms abstraits ; l'infixe *-an-*, des noms concrets, et le suffixe *-an*, des noms de localité :

ka- : magindanao *kaputi* « blancheur » < *puti* « blanc », tontemboan *kawĕlar* « largeur » < *wĕlar* « étendre », dayak *kagogop* « soin » < *gogop* « dont il est pris soin », sunda *kañaho* « connaissance », toba-batak *halinu* « réflexion (d'une image) » < vieux-bugi *lino* « se réfléchir », merina *hatsára* « bonté » < *tsára* « bon ».

-an- : bisaya *tanoptop* « son lointain » < *tup*, bugi *kanuku* « griffe » < *kuku*, madura *sanolap* « jonglerie » < javanais *sulap* « jongler », toba-batak *hanapa* « fibres qui enveloppent » < vieux-javanais *kapa* « enveloppe », formose *kaloṅkoṅ* < *kanoṅkoṅ* [1] « griffe » < ilokan *koṅkoṅ* « égratigner », kamber *tanai* « intestins » < indonésien commun *tai* « excrément », mentaway *tanai* « excrément » < *tai*.

-an : magindanao *niugan* « forêt de cocos » < *niug* « coco-

1. Pour l'alternance *n* > *l* en formosan, voir ci-dessus, p. 418.

tier », bugi *labuwañ* « ancrage », dayak *kayuan* « forêt », sunda *tandzakan* « terrain montant », toba-batak *hundulan* « endroit où l'on s'assoit », malais *babuhan* « ancrage », merina *sampánana* « ramification du chemin » < *sámpana* « se ramifier », simalur *kubañan* « mare à buffles » < malais *kubañ* « se vautrer dans la boue ».

De plus, chaque série de verbes différenciés par leur morphème verbal, possède un nom d'agent correspondant : malais *mĕmbunuh* « tuer », *pĕmbunuh* « assassin » ; malgache *mamunu* « tuer », *mpamunu* « assassin ». Dans les deux cas, la finale des préfixes verbaux *mĕñ-*, *man* et nominaux *pĕñ-*, *mpan-* est passée à la nasale labiale devant les labiales *b* et *v* des thèmes radicaux malais *bunuh* et malgache *vunu*, conformément à la loi de sandhi qui régit la formation des complexes.

Dans certaines langues indonésiennes, quand on fait le décompte d'êtres humains, d'animaux ou d'objets, les noms simples ou thèmes radicaux nominaux sont répartis en classes caractérisées par un déterminatif spécial. Ces déterminatifs sont les suivants en malais, où ils sont représentés en plus grand nombre : *orañ* « personne » pour les êtres humains et les anges ; *ikor* « queue » pour les animaux, même quand ils n'ont pas de queue, comme les grenouilles et les mouches ; *buah* « fruit », pour les fruits, les maisons, les villes, les navires, les îles, les lacs ; *bidzi* « graine », pour les graines et les petits objets plus ou moins ronds ; *batañ* « tige, tronc » pour les objets longs ; *kĕpiñ* « plat » pour les objets plats (planches, pièces de monnaie, étoffes, feuilles d'arbres), etc. Ainsi : *buda' dua orañ* « deux enfants, litt. enfants deux personnes », *kuda tiga ikor* « trois chevaux, litt. chevaux trois queues », *rumah dua buah* « deux maisons, litt. maisons deux fruits ». J'ai retrouvé quelques exemples de déterminatifs dans des textes malgaches, mais la langue moderne ne les connaît plus.

De deux noms en apposition, le second est le complément du premier : malais *astana radza* « [le] palais [du] roi », javanais *ratu dzawa* « [le] roi [de] Java », malgache *ela-papañgu* < *elatra* + *papañgu* « aile [de] milan » ; malais *rumah batu* « maison [en]

pierre », malgache *vilani vi* « marmite [en] fer ». Dans une phrase nominale telle que malais : *hikayat pĕrmula'an kadzadi'an sĕgala makluk*, les trois derniers noms sont compléments de celui auquel ils sont postposés : « histoire [du] commencement [de la] création [de] les (*sĕgala*) créatures ». La construction inverse du type anglais *birthday*, allemand *Haustor*, n'est usitée que dans quelques langues (rotti, solor, timor, flores, leti, kisar, bañgai, sula). Mais la règle précédente n'est pas absolue : le nom complément d'un autre nom est souvent régi par une préposition : dans vieux-javanais *tanah ri Malayu* « le pays de Malayu », malgache *tranu-m-buruna* (*tranu* + *n* + *vuruna*) « nid de [l']oiseau », les prépositions indonésiennes *ri* et *n* répondent exactement au *de* français.

L'ARTICLE. Il en est sept sortes :

1° l'article personnel : *ra, sañ, si, i, on* > *an* ;

2° l'article de parenté : *ša*, en mentaway ;

3° l'article défini : *ni, na, nan, ñi, ñi, añ, a, e* ;

4° l'article indéfini : *esa, mesa,* litt. « un » ;

5° l'article collectif : *para*, en javanais : *para pradzurit* « les soldats », litt. l'ensemble des soldats ;

6° l'article-ligature : *ñ, ña, ai, i, ka* ;

7° l'article partitif : *nan, nañ.*

Ra en vieux-javanais et en malgache a un sens respectueux et ne s'emploie qu'avec les noms divins et royaux : vieux-javanais *ra hyañ* « la divinité » ; malgache *Ra Dama* « Monseigneur ou Sa Majesté Dama » (nom de roi). C'est également le cas de *sañ* en vieux-javanais. *Ra* en malgache et *si* en atšeh s'emploient ironiquement dans des expressions telles que : malgache *Ra vwalavu* « Monseigneur le Rat », atšeh *Si buñkuk* « Monseigneur le Bossu ». *Si*, en malgache, est usité devant des noms communs de choses inanimées en fonction de noms propres géographiques : *Si hanaka* « Monseigneur le lac » (territoire habité par les Antaisihanaka « les gens [qui habitent autour] du lac [Alautra] ». *On* > *an* est, en malgache oriental, l'article des noms tribaux : *on-Tañala*, en malgache oriental ancien ; *an-Tañala*, en mal-

gache oriental moderne « les Tañala » (litt. *on* > *àn* « les » ; *ta* « gens », *añ ala* « dans la forêt », « les gens qui vivent, qui habitent en forêt »). Dans les dialectes malgaches orientaux modernes, *an* est passé à *ā*, de même que *on* devait être passé à *ō* au XVIIᵉ siècle où il nous est attesté ainsi par des relations de voyages et des textes arabico-malgaches de cette époque.

A Madagascar, l'article défini *ni* du merina est représenté par *ñi* ou *ṅi* dans les dialectes de l'Est et de l'Ouest.

Parmi les groupes suivants qui comprennent des langues étroitement apparentées l'une à l'autre, on constate les concordances et divergences suivantes.

En atšeh et gayo du Nord de Sumatra, l'article défini et l'article indéfini font défaut. *Si* est article personnel en atšeh et pronom en gayo. Le démonstratif *a* du gayo souvent employé en fonction d'article, manque en atšeh.

Dans le groupe batak du Nord de Sumatra, ni le karo ni le toba ne possèdent l'article défini ni l'article indéfini ; mais ils ont tous deux l'article personnel *si*. Toba et karo emploient fréquemment le démonstratif en fonction d'article : *i*, *on*, *inon* en toba ; *ndai* en karo.

En bĕsĕmah, sĕraway, minaṅkabaw et lampoṅ de Sumatra et malais de Malaka, l'article défini et l'article indéfini font défaut. L'article personnel est représenté en malais par *si* et *saṅ* ; en minaṅkabaw par *si* ; dans les autres langues par *saṅ*. Les démonstratifs *sena* en lampoṅ, *itu* dans les autres langues sont souvent usités en fonction d'article.

En madura et kaṅean, l'article défini et l'article indéfini manquent. Les deux langues possèdent les articles personnels *se*, *saṅ*. En javanais moderne, sunda et bali, l'article défini *ṅ* n'existe que dans quelques formules, et il y a lieu de croire qu'il s'agit d'une survivance du vieux-javanais. L'article personnel est *i* à Bali, *si* dans les autres langues.

A Rotti et Kupaṅ, le nom de nombre « un » est usité en fonction d'article indéfini. L'article défini est à Rotti : sing. *a*, plur. *la* après une voyelle, *ala* après une consonne ; à Kupaṅ : sing. *lia*, *la*, *li* ; plur. *lias*, *las*, *lis*. Dans ces langues et quelques

autres, l'article se place après le nom : Rotti *beis a*, litt. « crocodile le », *kaboa makadotok ala*, litt. « fruits *kaboa* nombreux les ».

A Bornéo, les dialectes tarakan et boloñan du groupe tidoñ ne possèdent ni l'article défini, ni l'article indéfini, mais offrent quelques exemples d'emploi des articles personnels *i* et *si*. Le dialecte boloñan emploie souvent en fonction d'article le démonstratif *inan* ; le dialecte tarakan, le démonstratif *gina*. Le dayak connaît l'article personnel *i* pour les noms propres et les titres officiels et emploie souvent en fonction d'article le démonstratif *dẑetä*, forme réduite *tä*.

Dans le Sud de l'île de Célèbes, le makassar et le bugi ont l'article défini *a* dans le premier dialecte, *e* dans le second, qui se place après le nom : *añiñ a*, litt. « vent le ». Le makassar connaît deux articles personnels *si* et *puñ*, celui-ci dans la langue poétique ; le bugi également : *la* pour les hommes, *i* pour les femmes.

Dans le Nord de l'île de Célèbes, l'article personnel est *si* en tontemboan, *ki* en bolaañ-moñondow. L'article défini est *ĕn* en tontemboan, *in* en bolaañ-moñondow.

Aux Philippines, le tagal et le bisaya emploient également l'article personnel *si* et l'article indéfini *añ*. Le tagal connaît l'article personnel pluriel *sina* : *sina Juan*, litt. « les Jean », c'est-à-dire « Jean et les siens, les personnes qui l'accompagnent ».

L'article-ligature (ce nom de *ligature* est emprunté aux grammairiens espagnols des Philippines) est connu du vieux-javanais : *ika ñ anak* « l'enfant, litt. celui-ci le enfant » ; mais il est plus spécialement employé dans les langues et dialectes des Philippines, ainsi tagal : *aso ñ malaki*, littéralement « chien le grand, un grand chien » ; *añ aso ñ malaki*, littéralement « le chien le grand, le grand chien ».

Les noms de nombre. On n'indiquera ici que les dix premiers dans les langues parlées à la périphérie du domaine, en tontemboan et en bugi. L'indonésien emploie la numération décimale.

Indonésien commun	Puyuma	Sumba	Mentaway	Merina	Tontemboan	Bugi
1 *sa	sa	sa	sa	isa	ĕsa	si
2 *dua	rua	dua	rua	rua	rua	duwa
3 *tĕlu	tero	tilu	tälu	telu	tĕlu	tĕllu
4 *ĕpat	spat	patu	äpat	efatra	ĕpat	ĕppa'
5 *lima	rima	lima	lima	dimi	lima	lima
6 *ĕnĕm	unum	nomu	änäm	enina	ĕnĕm	ĕnnĕn
7 *pitu	pitu	pitu	pitu	fitu	pitu	pitu
8 *walu	waro	walu	balu	valu	walu	aruwā
9 *siwa	iwa	siwa	šiba	sivi	siow	aserā
10 *puluh	purru	kĕmbuluh	pulu	fulu	pulu'	pulo

Les pronoms personnels ont une double forme, selon qu'ils sont à l'état isolé ou en fonction de pronoms suffixes. Quelques langues ont un pronom sujet et un pronom régime. Par exemple, malgache *aho* « je », *ahi* « me » ; *izi* « il », *azi* « lui ». Le pronom de la 1re personne du pluriel est double : l'un, dit inclusif, désigne la ou les personnes à qui l'on parle ; l'autre, dit exclusif, s'applique à la généralité : malais, *kita*, malgache *isika* « nous (ici présents dans cette salle) » ; malais *kami*, malgache *izahay* « nous (les Malais, nous faisons telle chose) ».

L'indonésien ne connaît pas le duel [1].

Doublets bilingues. C'est une curieuse pratique, usitée en indonésien occidental, qui méritait d'être signalée. Voici les exemples connus de doublets bilingues :

Sanskrit-kavi : *çapatha maṅmaṅ* « malédiction » ; sanskrit-sunda *candra-wulan* « Lune », nom d'une princesse ; javanais-sanskrit *ular-nāga* « serpent » ; sanskrit-malais *kandra-kayet* « crochet » ; persan-malais *dżawhar-manikam* « joyau », nom de l'héroïne princière d'un roman malais ainsi appelé (le second terme du complexe est un mot malais emprunté au sanskrit).

La réduplication du thème radical est fréquente. Elle exprime que l'état ou l'action exprimés par le thème est renforcé, dimi-

1. Voir ci-dessus, p. 414, pour le pronom personnel neutre du groupe paulohi.

nué, amélioré ou rendu pire ; que l'action est renouvelée fréquemment, etc. ; vieux-javanais *añin* « vent », *añin-añin* « tempête » ; tontemboan *lĕñei e lĕñei* « très pauvre » ; malgache *futsi* « blanc », *futsi-futsi* « blanchâtre ».

La réduplication se présente quelquefois sous la forme de la multiplication du thème radical. Ainsi, en mentaway, *igi* « nombreux » est répété quatre fois dans l'expression *igi-igi-igi-igi* avec le sens de « plus nombreux que tout ce qu'on peut imaginer ».

La phrase active avec l'ordre : sujet + verbe actif + complément se trouve dans toutes les langues du domaine. Mais la construction passive est plus fréquente ; c'est le mode favori d'expression de la pensée indonésienne. Une phrase telle que : « le fonctionnaire trancha la tête du Fakir », peut se rendre en suivant l'ordre des mots dans toutes les langues indonésiennes ; un lettré indigène en trouvera, cependant, l'expression un peu plate. Mais le passif semble plus élégant ; ainsi dans le texte bugi *Budi-Isĕtiharatĕ : na-ripolo-na ĕllon-na pakkerĕ'-e ri-to-ma'radža-e*, litt. « il (neutre) fut tranché tête-sienne Fakir-du par fonctionnaire-le = il fut tranché la tête du Fakir par le fonctionnaire (litt. le grand homme) ».

Écriture.

Les systèmes graphiques anciennement et actuellement usités dans les îles de l'Indonésie sont de trois sortes :

1° l'alphabet sanskrit et les alphabets dérivés d'alphabets de l'Inde. Celui-là a été utilisé dès le ve ou peut-être le ive siècle de notre ère dans l'épigraphie. Les anciens alphabets philippins d'origine indienne n'étaient déjà plus en usage en 1745 (voir *A pronouncing gazetteer and geographical dictionary of the Philippine islands*, 1902, p. 69). En Indonésie, les indigènes utilisent des alphabets d'origine indienne à Célèbes (bugi et makassar), à Ende ou Flores et Bima, à Bali, Madura, Java (kavi ou vieux javanais et javanais moderne), à Sumatra (batak, redžañ, rentšoñ, lampoñ).

2° l'alphabet arabe, augmenté des phonèmes indonésiens qui

n'existent pas en arabe, est en usage chez les Malais de la péninsule, à Sumatra (malais, atšeh minańkabaw), à Ternate des Moluques et chez les Moros ou musulmans des Philippines.

On utilise encore l'alphabet arabe pour la transcription du javanais et du sundanais (Ouest de Java); ce dernier système graphique est appelé *pegon*.

3° l'alphabet latin a été répandu et se répand partout ailleurs grâce aux développements des écoles hollandaises, officielles et confessionnelles.

BIBLIOGRAPHIE

J. L. A. Brandes. *Bijdrage tot de vergelijkende klankleer der westersche afdeeling van de Maleisch-Polynesische taalfamilie*. Utrecht, 1884, in-8.

Renward Brandstetter. *Malaio-polynesische Forschungen*. Lucerne :

Monographies in-4° : *Der Natursinn in den älteren Litteraturwerken der Malaien*, 1893; *Die Beziehungen des Malagasy zum Malaiischen*, 1893; *Die Geschichte von Hang Tuwah (Ein älterer malaiischer Sittenroman)*, 1894; *Die Geschichte von König Indjilai (Eine bugische Erzählung)*, 1895; *Die Gründung von Wadjo (Paupau Rikadong), eine historische Sage aus Südwest-Selebes*, 1896; *Drei Abhandlungen über das Lehnwort (Das Lehnwort in Luzerner Mundart, Das Lehnwort in der bugischen Sprache, Die Lehnwörter, welche der Luzerner Mundart und der bugischen Sprache gemeinsam angehören)*, 1900.

Monographies in-8° :

Die Geschichte von Djajalankara (ein makassarischer Roman), 1898; *Tagalen und Madagassen (Eine sprachvergleichende Darstellung als Orientirung für Ethnographen und Sprachforscher)*, 1902; *Ein Prodromus zu einem vergleichenden Wörterbuch der malaio-polynesischen Sprachen für Sprachforscher und Ethnographen*, 1906; *Mata-Hari oder Wanderungen eines indonesischen Sprachforschers durch die drei Reiche des Natur*, 1908; *Wurzel und Wort in den Indonesischen Sprachen*, 1910; *Sprachvergleichendes Charakterbild eines Indonesischen Idiomes* (le bugi de Célèbes), 1911; *Gemeinindonesisch und Urindonesisch*, 1911; *Das Verbum dargestellt auf Grund einer Analyse der besten Texte in vierundzwanzig indonesischen Sprachen*, 1912; *Der Artikel des Indonesischen verglichen mit dem des Indogermanischen*, 1913; *Indonesisch und Indogermanisch im Satzbau*, 1914; *Die Lauterscheinungen in den Indonesischen Sprachen*, 1915; *Die Reduplikation in den indianischen, indonesischen und indogermanischen Sprachen*, 1917; *Architektonische Sprachverwandtschaft in allen Erdteilen*, 1920; *Wir Menschen der indonesischen Erde*. I. *Die indonesische und die indogermanische Volksseele, Eine Parallele auf Grund sprachlicher Forschung*, 1921.

Encyclopaedie van Nederlandsch-Indië, 4 vol. in-4°, 1re édit.

GABRIEL FERRAND. *Essai de phonétique comparée du malais et des dialectes malgaches*. Paris, 1909, in-8°.

A. A. FOKKER. *Malay phonetics*. Leyde, 1895, in-8°.

J. C. G. JONKER. *Kan men bij de talen van den Indischen archipel eene westelijke en eene oostelijke afdeeling onderscheiden ?* Verslag. en Meded. der K. Akademie van Weten., afdeeling Letterkunde, 4e Reeks, deel XII, 1914, p. 314-417 ; *Rottineesch-hollandsch woordenboek*. Leyde, 1908, in-4° ; *Rottineesche spraakkunst*. Leyde, 1915, in-4°.

H. KERN. *Verspreide geschriften* (depuis 1913, 10 volumes ont paru).

Abbé ROUSSELOT. *Phonétique malgache*, dans *Revue de Phonétique*. Paris, 1913, 103 p.

ALB. T. SCHWARZ. *Tontemboansch-nederlandsch woordenboek*. Leyde, 1908, in-4°, publié par le Dr N. ADRIANI ainsi que l'ouvrage suivant : *Hoofdstukken uit de spraakkunst van het tontemboansch* (en collaboration avec Mme ADRIANI). *Ibid*.

ERWIN STRESEMANN. *Die Paulohisprache, ein Beitrag zur Kenntniss der Amboinischen Sprachengruppe*. La Haye, 1918, in-8°.

On trouvera de précieux renseignements sur l'indonésien dans les deux revues suivantes :

Bijdragen tot de taal-, land- en volkenkunde van Nederlandsch-Indië, La Haye, qui paraît depuis 1853. Un index des 50 premiers volumes a été publié en 1901 ;

Tijdschrift voor indische taal-, land- en volkenkunde uitgegeven door het Bataviaasch genootschap van Kunsten en Wetenschappen, Batavia. Un index des volumes parus depuis le début a été publié en 1908.

Voir également les études comparatives de H. Kern mentionnées dans la bibliographie du mélanésien, et les *Wissenschaftliche Ergebnisse einer amtlicher Forschungsreise nach dem Bismarck-Archipel im Jahre 1908* ; t. II, *Beiträge zur Völker- und Sprachenkunde von Deutsch-Neuguinea*, par GEORG FREDERICI, Berlin, 1912, in-4° ; t. III, *Untersuchungen über eine melanesische Wanderstrasse*, 1913, par le même auteur (ces deux volumes contiennent d'importantes études de linguistique comparée et une bonne bibliographie du sujet).

Dans un mémoire intitulé : *Die Mon-Khmer-Völker, ein Bindeglied zwischen Völkern Zentralasiens und Austronesiens* (Braunschweig, 1906), le Père Schmidt a montré la parenté des langues malayo-polynésiennes avec les langues mon-khmer et par l'intermédiaire de celles-ci avec certaines langues de l'Inde. Ce mémoire a été traduit en français dans le *Bulletin de l'École Française d'Extrême-Orient*, t. VII, 1907. Voir ci-dessus, p. 385, une critique de cette théorie.

MÉLANÉSIEN

Répartition géographique.

Les îles de la Mélanésie s'étendent, en direction du Nord-Ouest au Sud-Est, de l'équateur jusque vers le 17ᵉ degré de latitude méridionale, entre les 140ᵉ et 180ᵉ degrés de longitude Est de Paris.

Les langues et dialectes dont il sera question sont répartis dans les îles suivantes, énumérées du Nord-Ouest au Sud-Est.

Archipel des Salomon :

Ile du duc d'York,

Ile Ysabel : Ṅao et Bugotu,

Ile Savo,

Ile Florida,

Ile Guadalcanar : Vaturaṅa,

Ile San Cristoval : Fagani et Waṅo,

Ile Malanta : Saa,

Ile Ulawa (Contrariété Island, *sic*).

Archipel de Santa Cruz :

Ile Deni et Nifilole,

Ile Rotuma, à l'Est des précédentes.

Archipel de Torres : Lo.

Archipel de Banks :

Ile Ureparapara,

Ile de la Selle (Saddle Island) : Volow et Motlav Mota (Sugarloaf Island),

Ile Vanua Lava : alo Tequel, Mosina, Vureas, Sasar, Pak.

Ile Santa Maria : Lakon, Gaua ou Gog.

Ile Merlav (Star Island).

Archipel des Nouvelles-Hébrides :

Ile Aurora : Maewo,

Ile des Lépreux (Lepers' Island) : Oba,
Ile Araga (Whitsuntide island ou île de la Pentecôte),
Ile Espiritu Santo,
Ile Ambrym,
Ile Sesake (Three Hills),
Ile Sandwich : Fate,
Ile Eromanga,
Ile Aneityum.
Archipel des Iles Loyalty :
Ile Mare ou Nengone,
Iles Fidji ou Fiji.

Pendant la colonisation allemande des îles du Nord-Ouest, celles-ci ont pris les nouveaux noms suivants : la Nouvelle-Bretagne est devenue Neu-Pommern et sa partie septentrionale, la presqu'île de la Gazelle ; l'île du duc d'York, Neu-Lauenburg ; la Nouvelle-Irlande, Neu-Mecklemburg ; la petite île au Nord-Ouest de cette dernière : Neu-Hannover. L'ensemble de ces îles a été dénommé archipel de Bismarck.

Les langues et dialectes mélanésiens ont été étudiés par le Rév. R. H. Codrington dans *The Melanesian languages* (Oxford, 1885, in-8°). On désignera les îles dont il s'agit sous les noms employés par ce savant. Pour la langue de Fidji, voir H. Kern, *De Fidjitaal vergeleken met hare verwanten in Indonesië en Polynesië* (*Verh. der k. Akad. d. Wet. Afd. Letterk.*, deel XVI, Amsterdam, 1886) et *Taalvergelijkende over het Aneityumsch, met een Aanhangsel over het klankstelsel van het Eromanga* (*Verh. der k. Akad. van Wet. Afd. Letterk.*, 9 reeks, deel VII, n° 2, Amsterdam, 1906).

Phonétique.

Le système phonétique du mélanésien se compose de :
Sept voyelles : *a, e, i, o, ö, u, ü.*
cinq diphtongues : *ay, ae, ao, au, aw.*
deux semi-voyelles : *y, w.*
cinq gutturales : *k, ṅk, g, ṅg, ṅ.*
cinq prépalatales : *tš, ttš, dž, ndž, ñ.*

une affriquée : *ts*.
quatre dentales : *t̨*, *d*, *nd*, *n*.
cinq labiales : *p*, *b*, *mb*, *m*, *m̤*.
deux liquides : *r*, *l*.
trois sifflantes : *s*, *z*, *ʒ* (sorte de *z*).
une sifflante interdentale : θ (*th* anglais de *thin*).
une aspirée : *h*.
deux dentilabiales : *f*, *v*.
quatre phonèmes spéciaux : *kw*, *pw*, *ṭ*, *ḍ*.

Les voyelles sont longues ou brèves à l'initiale et à la médiale ; à peine perceptibles à la finale.

Codrington fait remarquer (p. 199, note 3) que *ao*, *au*, *aw* représentent trois diphtongues nettement différenciées, par exemple, dans *gau* « hameçon », *gao* « se répandre (se dit du feu ou des nouvelles) » et *gaw* « prendre à poignée ».

On sait que les langues mélanésiennes n'ont jamais possédé de système graphique. Elles n'ont été notées que par les voyageurs européens et, en dernier lieu, par les missionnaires chrétiens qui se sont établis en Mélanésie, et ces notations sont loin d'être toujours satisfaisantes.

D'après les travaux de Codrington (*The Melanesian languages* et *A dictionary of the language of Mota, Sugarloaf Island, Banks' islands*, Londres, 1896, in-8°), quelques phonèmes sont précédés d'une nasale de leur classe : suivant les langues ou dialectes, les gutturales sourde et sonore sont usitées à l'état de gutturales pures, *k*, *g*, ou de gutturales prénasalisées : *ňk*, *ňg*. Il en est de même pour la dentale et la prépalatale sonores : *d* ou *nd*, *dž* ou *ndž*, et la labiale sonore : *b* ou *mb*.

L'alternance *tš* et *ttš* (Codrington écrit *tch* en orthographe anglaise), c'est-à-dire, dans le second cas, *t* + *tš* est moins claire.

Codrington (*Melanesian languages*, p. 198) note deux sortes d'*m* : un *m* ordinaire imprimé en romain et un *m* en italique avec l'indication expresse qu'il s'agit d'un *m* nasal (*sic*). L'auteur ajoute en note : « écrit aux îles Loyalty **m* et dans le sud des Nouvelles-Hébrides *mw* ». C'est cette sorte d'*m* qu'on a représenté ici par *m̤*. En mota, un grand nombre de mots, homophones par ailleurs,

ne sont différenciés que par l'un de ces deux *m* : *iṃa* « maison », *ima* « boire » ; *taṃa* « comme », *tama* « père » [1].

Les deux phonèmes *kw* et *pw* sont également rendus par *q* dans la transcription des missionnaires anglais. « C'est un composé, dit Codrington (*Melanesian languages*, p. 198), de *kpw* dans lequel quelquefois *p* n'apparaît pas, quelquefois *k* est à peine perceptible. Les lèvres sont fermées pour la formation de la gutturale et ouvertes un peu subitement pour émettre la phonation. Le son [ainsi obtenu] tend vers *kw* ou *pw* suivant que la gutturale ou la labiale est plus nettement articulée. »

Les deux phonèmes qu'on a transcrits *ṭ* et *ḍ* se différencient des dentales sourde et sonore en ce qu'ils laissent percevoir un élément vibrant qui les a fait rendre quelquefois par *tr* et *dr*. Il semble que nous ayons affaire ici à des phonèmes à peu près identiques aux *tr* et *dr* malgaches (voir plus haut, p. 416). Ainsi Fidji, *ḍau* (*drau*) « feuille d'arbre » ; tagal, bisaya *dahu*, bikol *dahon*, moṅondow, ponosakan *daun*, malais *dāun*, vieux-javanais *ron*, *rwan*, saṅir, sumba *rau(n)*, dayak *dawĕn*, malgache *rávina*, bugi *dauṅ*, makassar *rauṅ* ; Fidji *drau* « cent » < indonésien commun *ṛatus*. Parallèlement, *tr* et *dr* malgaches remontent généralement à *d*, *r*, *l* indonésiens.

La sifflante ʒ — c'est la transcription qu'a adoptée Kern dans sa *De Fidjitaal* — répond aux phonèmes indonésiens suivants : ʒ < *s* (Fidji *ʒiwa* « 9 » < indonésien ou plutôt malayo-polynésien commun *siwa*) ; ʒ < *ḍ*, *r*, *d*, *g* (Fidji *taʒi* « jeune frère ou jeune sœur » < malais *adi'*, vieux-javanais *ari*, javan. moderne

1. Dans sa *Grammar of the language of Sa'a, Malaita, Salomon islands* (*Anthropos*, t. VI, 1911, p. 755), le Rév. W. G. Ivens dit : « la prononciation de *ṃ* varie entre *m* nasal [*sic*], *mw* et *m* pur ». D'après le Père J. Rausch, « *ṃ* est légèrement murmuré. Il ne se présente qu'à l'initiale, principalement dans les mots qui indiquent une relation de parenté, comme *ṃa* « mon père », *ṃama* « ma sœur aînée », *ṃe* « pas encore », *ṃarui* « il s'est enfui ». (*Die Sprache von Südost-Bougainville, Deutsche Salomoninseln*, dans *Anthropos*, t. VII, 1912, p. 105.) Les deux auteurs précédents représentent ce *m* spécial par une *m* surponctué dont ni l'un ni l'autre n'indiquent le véritable timbre. Les langues dont il s'agit sont parlées dans des îles mélanésiennes, mais appartiennent, d'après les missionnaires Ivens et Rausch, à la famille linguistique papoue.

et sunda *aḍi*, lampon *adiụ*, makassar *ari*, *andi*, bugi *anri*, *andi*, dayak *andi*, batak *añgi*, ibanag *agit*, malgache ancien, *yandri*, malg. moderne *zandri*, nias *aḫi*, tombulu, tondan, sawu, alor *ari*, timor *oli* ; Fidji *viða* « combien ? » < (javanais *pic̣a*, batak et ibanag *piga*) ; ð < *dž* (Fidji *ða* « mauvais, méchant » < malais *džahat*), ð < *tš* (Fidji : préfixe *ða-* < bugi *tša-*).

Le tableau précédent des 32 consonnes du mélanésien ne doit pas faire illusion : il s'en faut de beaucoup qu'elles soit représentées dans toutes les langues du groupe, ni même dans une seule. Tout d'abord, cinq consonnes se présentent soit à l'état pur soit à l'état prénasalisé par une nasale de même classe. *K* et *ñk*, *d* et *nd*, *b* et *mb*, etc., sont de simples doublets. « Le *b* pur, dit Codrington (*Melanesian languages*, p. 209), par exemple, est à peine entendu en Mélanésie, en dehors de San Cristoval et des îles Loyalty. Il est difficile de le distinguer du *p*, d'une part, du *v*, de l'autre. A Araga, certaines gens prononcent *b*, d'autres *mb* ; et le même mot peut être prononcé indifféremment *pev*, *bev*, *mbev*, *vev* ». Les autres phonèmes se répartissent entre les langues du groupe d'après les lois d'alternances qu'il reste à fixer dans le détail. La phonétique du mélanésien n'a pas été étudiée encore de façon assez sûre pour qu'on puisse faire état des travaux que nous possédons à cet égard.

Morphologie.

a) Description sommaire de la langue des îles Fidji.

Nous connaissons assez bien la langue de la plus importante des îles de l'archipel de Fidji (*Fidji* ou *Fiji* est une notation européenne pour *Viti*, plus exactement *Viti levu* « la grande Fidji », qu'on a transcrit également *Fiti leb*, *Biti lib* et *Metalip* = *Mbiti levu*, variante phonétique de *Viti levu*).

Remarque phonétique. — « La langue parlée, dit Friedrich Müller (*Grundriss der Sprachwissenschaft*, t. II, 2e partie, Vienne, 1882, p. 51), ne différencie pas les sourdes *k*, *t* des sonores *g*, *d*. Les sonores *g*, *d*, *b* sont prénasalisées : *ñg*, *nd*, *mb* ; la vibrante *r* se prononce *ndr*. »

Le thème radical possède les deux sens nominal et verbal que précise seul le contexte : *vosa* signifie également « parler » et « discours » ; ainsi *au sa vosa* « je parle » ; *a vosa* « le discours ». La réduplication du thème radical, quel qu'en soit le sens, produit :

des verbes fréquentatifs : *kere* « demander », *kerekere* « mendier » ; *katsi* « appeler », *katsikatsi* « appeler fréquemment » ;

des noms de lieu, d'instrument, etc. : *ndave* « couler », *ndavendave* « canal » ; *mbulu* « enterrer », *mbulumbulu* « cimetière » ;

des adjectifs : *loaloa* « noir, sale » < *loa* « excrément, saleté » ; *ndrendre* « lourd » < *ndre* « tirer ».

Préfixes. Le préfixe *ndo-* forme : 1° des verbes fréquentatifs : *ndau-lolo* « jeûner souvent » ; 2° des noms d'agent : *ndau-mbutako* « voleur » (*mbutako* « voler ») ; 3° des adjectifs qui indiquent une propriété, qualité, manière d'être habituelle : *ndau-loloma* « miséricordieux » (celui qui a habituellement bon cœur).

Vei- forme : 1° des verbes réciproques : *vei-lomani* « s'aimer mutuellement » ; 2° des noms collectifs : *vei-vanua* « des pays », *vei-kau* « des arbres » ; 3° des adjectifs indiquant la plénitude : *vei-vatu* « pierreux, riche en pierres ».

Vaka- forme : 1° des verbes causatifs et potentiels (ainsi indonésien *paka*, *maka*, *maha*; polynésien *whaka*, *faa*) : *vaka-mbulu* « rendre vivant (*mbulu* « vivre, vivant ») ; 2° des adjectifs indiquant la similitude avec le thème radical ou la possession du thème radical auquel il est préfixé : *vaka-tamata* « semblable à l'homme », *vaka-vale* « possédant une maison ».

Les préfixes précédents s'emploient aussi simultanément dans un ordre déterminé : *ndau-vei*, *ndau-vei-vaka* et forment ainsi des verbes réciproques-fréquentatifs : *ndau-vei-moku* « se battre violemment mutuellement », *ndau-vei-vaka-tsundrui* « se mettre souvent mutuellement en colère ».

Ka-, *ta-*, *ra-* sont exclusivement préfixes verbaux et forment des verbes médio-passifs : *ka-ndresu* « se casser, être cassé « de

ndresu-ndresu « se casser » ; *ta-rambe* « se heurter » de *rambe* « heurter du pied » ; *ra-musu* « être cassé « de *musu* « couper ».

Suffixes. Les suffixes sont monosyllabiques : *-a*, *-ka*, *-ñga*, *-ta*, *-tsa*, *-ya*, *-ra*, *-va*, *-wa*, *-na*, *-ma*, ou dissyllabiques : *-kaka*, *-taka*, *-tsaka*, *-laka*, *raka*, *vaka*, *waka* et forment exclusivement des verbes transitifs.

Pronoms personnels. Comme l'indonésien, la langue de Fidji possède une série de pronoms isolés et une série de pronoms suffixes, avec un double pronom, isolé et suffixé à la 1re personne du pluriel : l'inclusif et l'exclusif.

Le fidjien et les autres langues mélanésiennes possèdent, en outre, le duel et le triel.

Verbe. Le paradigme du verbe comprend quatre temps : un aoriste-présent, un présent défini, un parfait et un futur qui s'expriment respectivement ainsi : *au lako* « je vais » ; *au sa lako* ; *au a lako* ; *au na lako*.

Noms de nombre. A Fidji, la numération est décimale et nettement indonésienne, à l'exception du nombre 10 :

1	*ndua*	6	*ono*
2	*rua*	7	*vitu*
3	*tolu*	8	*walu*
4	*va*	9	*tsiwa*
5	*lima*	10	*tini*

Article. Deux sortes sont usitées : l'article des noms communs et celui des noms propres. Ces articles sont différents suivant que le nom est au cas sujet ou au cas régime :

article des noms communs : *na* (forme réduite *a*) au nominatif ; *ni* au génitif ; *ki na* au datif ; *ma na* à l'ablatif ; *i na* au locatif ; *kei na* à l'instrumental.

article des noms propres : *ko* (forme réduite *o*) au nominatif ; *i* au génitif ; *vei* ou *ki vei* au datif ; *mai* à l'ablatif ; *kinii* au locatif ; *kei* à l'instrumental.

Le nom à l'accusatif ne prend pas d'article et se place entre le verbe et le sujet qui est rejeté à la fin de la phrase : *sa kunei Natsanieli ko Filipe*, litt. « on trouva Nathaniel Philippe », c'est-à-dire Philippe trouva Nathaniel ».

b) Indications sur les langues mélanésiennes en général

On a vu précédemment (p. 426-7) que le thème radical malais *tidor* a le double sens de « dormir » et « sommeil » ; ce fait linguistique, qui est extrêmement rare en indonésien, est, au contraire, relativement commun en mélanésien. Ainsi Fidji *mate* « mourir » et « mort », et ci-dessus, p. 441.

Noms. Les différentes classes de noms sont les suivantes :

1° noms d'état ou d'action formés par la suffixation au thème radical, généralement verbal, de morphèmes *ana*, *na*, *ana*, *ena*, *ina*, *na* ;

2° noms formés par la suffixation du morphème *a* au thème radical verbal (Florida et Ysabel) ;

3° noms abstraits formés par la suffixation des morphèmes *a*, *ia*, *ea*, *e*, *va*, *v*, *ga*, *g*, *ra*, *r* (archipel des Banks et îles méridionales de l'archipel des Salomon), *be* (Saa des Salomon) ;

4° noms formés par suffixation des morphèmes *i*, *ge*, *gi*, *iu*, *ui*, *u*. Cette classe ne comprend que les noms désignant les parties d'un tout, comme les membres du corps, par exemple.

5° noms formés par préfixation du morphème *i-* à un verbe : fidji *sele* « couper », *i-sele* « couteau » ; mota *ras* « écoper l'eau d'une pirogue », *i-ras* « un écopeur ». Aux îles Banks, le morphème *ga-* a un rôle morphologique identique : *sal* « couper », *ga-sal* « couteau ».

Noms collectifs. Certains noms, simples ou redoublés, désignent un nombre déterminé d'objets : fidji *a uduudu* « 10 pirogues », *a buru* « dix cocos », *a bolas* « cent pirogues », *a selavo* « 1000 cocos » ; Florida *na gobi* « 10 pirogues », *na pigu* « 10 cocos », *na kua* « 10 œufs », *na pana* « 10 porcs, oiseaux, poissons ».

Réduplication. La réduplication totale ou partielle d'un mot marque l'intensité, la grandeur, le nombre, et aussi la diminution de taille ou de grandeur de la forme simple et la péjoration. Dans l'archipel des Salomon et des îles Banks, la réduplication du nom d'un fruit cultivé indique qu'il s'agit de l'espèce sauvage.

Les noms se divisent, en outre, en deux classes : ceux qui prennent le pronom possessif suffixe et ceux qui ne le prennent pas. « La distinction entre ces deux sortes de noms, dit Codrington (*Melanesian languages*, p. 143), est basée sur la proximité ou l'éloignement de la relation entre l'objet possédé et son possesseur ». Mais le même auteur ajoute que cette explication ne vaut pas pour tout le domaine et il en donne cet exemple décisif : mota *us* « arc », *na us-uk* « litt. le arc mien, mon arc (avec possessit suffixe) » ; motlav *ih* « arc », *n-ih mino* « litt. l'arc [de] moi (avec pronom personnel isolé) » ; et on n'aperçoit pas les raisons pour lesquelles « arc » est dans la première classe en mota et dans la seconde en motlav, alors surtout qu'il s'agit de deux îles voisines. En réalité, la classification précédente existe à n'en pas douter, mais on n'a pas su encore en retrouver l'origine ni le véritable emploi.

Le complément du nom se postpose au nom qui le régit : mota *o iṃa vatu* « la maison en pierre », fidji *a su ika* « le panier de poisson », sans préposition. L'emploi d'une préposition donne à l'expression un sens spécial : fidji *a su ni ika* « le panier pour (*ni*) le poisson ». En mota *iṃa kwoe* « la maison pour porc » est la porcherie, la maison exclusivement construite pour les porcs ; *iṃe kwoe* par changement en *e* de la voyelle finale *a* de *iṃa*, a, au contraire, le sens de « la maison du porc » (d'un porc déterminé).

Adjectifs. L'adjectif mélanésien se forme en suffixant au thème radical l'un des suffixes suivants :

Nouvelles-Hébrides : *-a*, *-ga*, *-gi*, *-sa* ;
Iles Banks : *-g*, *-r*, *-a*, *-ga*, *-ra*, *-la* ;
Fidji : *-a*, *-ta*, *-li* ;
Archipel des Salomon : *-a*, *-ga*, *-ha*, *-nia*.

Pronoms personnels. Le mélanésien possède, comme l'indonésien, deux séries de pronoms personnels : les pronoms isolés et les pronoms suffixes, avec un double pronom à la 1re pers. du pluriel, inclusif et exclusif. Il possède, en outre, un duel et un triel (inclusit et exclusif à la 1re personne) qui s'expriment par le pronom de la personne augmenté de *rua* « deux », *tolu* « trois ».

Verbes. Le verbe simple est transitif ou intransitif. Quatre séries de préfixe, donnent au verbe simple, le sens :

1° causatif (*va-*, variantes *wa-*, *ua-* ; forme secondaire *vaka-*, *e-*) ; *vaga-*, etc.) ;

2° réciproque (*vei-* ou *var-*, variantes *vui-*, *hei-*, *hai-*, *fai-*, *we-*, *e-*) ;

3° d'être dans l'état indiqué par le thème radical (*-ma-*, *-ta-*) ;

4° de spontanéité (*tava-*, *tapa-*) indiquant que l'action s'est accomplie automatiquement : mota *ul-* « dénouer (une corde nouée) », *ta-va-ul* « se dénouer spontanément, automatiquement », *me tava-ul* « (corde qui) s'est dénouée elle-même ».

Les temps du verbe s'expriment à l'aide de « particules verbales », pour employer la définition de Codrington, qui marquent : le présent indéfini, le passé, le futur, la continuité de l'action ou la permanence de l'état. Elles se placent devant le verbe.

Le mélanésien possède un très grand nombre de suffixes verbaux — en mota, par exemple, on en compte 17 : *-v*, *-r*, *-t*, *-s*, *-n*, *-ñ*, *-g*, *-ag*, *-gag*, *-lag*, *-ṃag*, *-nag*, *-ñag*, *-rag*, *-sag*, *-tag*, *-vag* ; 20 en fidjien — dont les fonctions n'apparaissent pas clairement, de l'aveu même des missionnaires qui les ont étudiés (*Melanesian languages*, p. 179). Ainsi, les verbes simples mota *roño*, fidjien *rogo* signifient « entendre » (intransitif) ; mais mota *roño-tag* a le sens de « entendre » (transitif) et fidjien *rogo-taka* celui de « dire, rapporter (une nouvelle) ». D'autre part, à l'intérieur du mota, deux suffixes différents donnent au verbe un sens identique : *saro-nag*, *saro-ṃag* signifie également « entrer ».

Contrairement à l'indonésien (ci-dessus, p. 433), le mélanésien n'use pas de la voix passive. Codrington déclare n'en

pas connaître d'exemple décisif (*Melanesian languages*, p. 191-192).

Article. Le mélanésien connaît deux sortes d'articles : l'article défini et l'article personnel. Celui-là est représenté par *n* + voyelle (généralement *na*) ; celui-ci par *i*, *e*, *a*.

Noms de nombre. La numération mélanésienne est quinaire dans certaines îles, décimale dans d'autres, vigésimale dans quelques langues. La numération décimale est indonésienne avec les alternances phonétiques attendues. Dans la numération quinaire développée jusqu'à 10, 6 s'exprime tantôt par 3 + 3, tantôt par 5 + 1 ; 7, par 5 + 2 et 6 + 1 ; 8, par 5 + 3 et 10 — 2 ; 9, par 5 + 4, 8 + 1 et 10 — 1 ; 10, par 2 × 5. A Fate des Nouvelles-Hébrides, 10 est *relima*, litt. 2 (fois) 5, et 20, *relima rua*, 2 (fois) 5 (multiplié par) 2. Aux îles Loyalty où la numération est quinaire, 20 se dit « un homme » (c'est-à-dire 4 fois cinq doigts des mains et des pieds) ; 30 = un homme (20) + 2 séries de cinq doigts ; 40 = deux hommes, etc.

BIBLIOGRAPHIE

Aux ouvrages déjà cités, ajouter :

R. Parkinson, *Dreissig Jahre in der Südsee*. Stuttgart, 1907, in-4°, p. 721-787.

MICRONÉSIEN

Les îles de la Micronésie gisent dans la partie du Pacifique comprise entre l'équateur et 20° Nord, d'une part ; et le 130e et le 176e degré de longitude orientale de Paris, d'autre part.

Elles comprennent :

Archipel des Gilbert ou Kingsmill Islands,

Archipel des Marshall,

Archipel des Carolines : Kusaie, Ponape, les îles Mortlock, Ṛuk,

Ile de Yap,

Archipel des Mariannes.

Les langues et dialectes des îles Palaw (Palau, Pelew) et le Chamorro de l'île Saipan des Mariannes font linguistiquement partie du groupe des Philippines et se rattachent ainsi à l'indonésien.

La phonétique du micronésien est, dans l'ensemble, identique à celle du mélanésien. Celle-là possède, en outre, une sifflante palatale : *ś*. L'*r* des îles de Ṛuk, Mortlock et de l'archipel des Mariannes a été transcrit tantôt *tr* ou *dr*, tantôt *t* ou *r*. Il s'agit en réalité, d'un phonème connu du mélanésien, qu'on a précédemment rendu (voir ci-dessus, p. 439) par *ṭ* ou *ḍ*. Il est noté ici *ṛ* pour le différencier de la vibrante linguale ordinaire.

Pour la morphologie et la syntaxe, le micronésien est assez proche du mélanésien et il n'y a pas lieu d'y insister. La catégorie pronominale est seule divergente et présente des particularités spéciales à ce groupe linguistique, uniques dans la famille malayo-polynésienne, qu'il y a lieu de mettre en lumière.

Le micronésien possède un duel, un triel et un quatriel, avec une forme inclusive et une forme exclusive pour la 1re personne.

Les noms auxquels se suffixent les pronoms possessifs sont divisés en classes :

1re classe : noms des parties du corps et des manifestations physiques, intellectuelles et morales de l'homme (âme, caractère, force, faiblesse, maladie, volonté, voix, etc.);

2e classe : noms de parenté ;

3e classe : prépositions nominales de temps et de lieu ;

4e classe : les mots « nom », « prénom », « surnom » ont un suffixe pronominal possessif spécial.

5e classe : noms de la partie d'un tout ou du tout en relation avec ses parties (racine, feuille, fruit d'un arbre ; ombre, image d'un corps, considérées comme émanation du corps ; noyau d'un fruit).

6e classe : noms de vêtements, parures, bijoux ; outils, ustensiles, instruments ; maison, habitation appartenant à quelqu'un ou à soi-même.

7e classe : noms possessifs, c'est-à-dire indiquant que telle ou telle chose est possédée par la 1re, 2e ou 3e personne du singulier ou du pluriel, que marque le pronom possessif.

8e classe : noms des comestibles solides.

9e classe : noms des comestibles liquides.

10e classe : elle n'existe que dans l'île de Kusaie et comprend des noms variés dont chaque catégorie a son possessif spécial : noms désignant les parties d'un village, d'une maison ; noms abstraits tels que « amour », « volonté », « joie » ; la préposition « à, vers » indiquant la direction (ainsi *ā siyuk nu se-l* « ils l'implorèrent » littéralement « ils implorèrent vers endroit (*se*) sien (*-l*) » ; etc.

M. A. Thalheimer a traité le sujet dans son *Beitrag zur Kenntniss der pronomina personalia und possessiva der Sprachen Mikronesiens* (Stuttgart, 1908, in-8°, avec une très complète bibliographie). L'étude est intéressante, mais il en ressort surtout, comme l'indique du reste l'auteur, que la question doit être reprise et

étudiée à nouveau. Les résultats qu'on peut considérer comme acquis sont importants, car ils établissent l'emploi en Micronésie d'un mode d'expression inconnu par ailleurs dans la famille malayo-polynésienne et qui n'a d'équivalent que chez des peuples de race et de langue entièrement différentes.

BIBLIOGRAPHIE

Voir ci-dessus, les ouvrages indiqués pour le mélanésien. Consulter également F. W. CHRISTIAN, *The Caroline islands*. Londres, 1899, in-8°, p. 324 et suivantes.

POLYNÉSIEN

Généralités.

La Polynésie comprend toutes les îles situées à l'Est et au Sud de la Mélanésie :

Iles Samoa,
Archipel de Cook ou îles Hervey : Rarotonga,
Iles de la Société : Tahiti,
Iles Paumotu ou Tuamotu,
Iles Tonga, Tongatabu,
Iles Gambier ou Mangareva,
Ile de Pâques,
Nouvelle-Zélande (Maori),
Iles Chatham ou Moriori.

Au Nord des îles précédentes : l'archipel des Marquises avec Nukahiva, par environ 10° Sud, et l'archipel des Hawaii, par environ 20° Nord. On parle, en outre, polynésien dans un certain nombre d'îles qui font géographiquement partie de la Mélanésie, telles que Uea, l'une des îles de l'archipel des Loyalty ; Futuna des Nouvelles-Hébrides ; quelques îlots du groupe Sheppard des Nouvelles-Hébrides et particulièrement Mae dans l'île Three Hills de cet archipel ; Tikopia, plusieurs îles du Swallow Group de l'archipel de Santa Cruz ; Rennel et Bellona au Sud de l'archipel des Salomon ; Ontoǹ Java, près de l'île Ysabel du même archipel (voir Codrington, *Melanesian languages*, p. 7-8).

Dans l'ensemble, le polynésien, qui est apparenté de très près à l'indonésien, réduit ou a une tendance à réduire à ses seuls éléments vocaliques le thème initial indonésien. Par exemple, malais *akar* « racine » > maori *aka* > tahiti et hawaii *aa* ; malais *ikan* « poisson » > marquises et mangareva *ika* >

samoa *i'a* > tahiti et hawaii *ia*. Cette loi du moindre effort ne s'applique pas seulement à la langue : les Polynésiens eux-mêmes ne réagissent pas contre l'excédent des décès sur les naissances et disparaissent rapidement ; tous les voyageurs le constatent. Ce peuple de belle allure, d'apparence robuste, a ses jours comptés sans qu'on sache exactement quelles sont les causes de son extinction prochaine.

Phonétique.

Le système phonétique du polynésien se compose des voyelles et consonnes suivantes :

a, e, i, o, u.

semi-voyelle : *w*
trois gutturales : *k, g, ṅ*
deux prépalatales : *tš, dž*
trois dentales : *t, d, n*
trois labiales : *p, b, m*
deux liquides : *r, l*
deux dentilabiales : *f, v*
une sifflante : *s*
une aspirée : *h*
une occlusive glottale : '

k manque à Samoa, Tahiti et Hawaii où il est passé à l'occlusive glottale. *t* s'est maintenu dans presque toutes les langues du groupe ; il passe quelquefois à *k* (Hawaii, Marquises) et plus rarement à *n* (Marquises). *p* s'est maintenu dans le groupe. *h* répond dans les autres dialectes à l'*s* de Samoa et de Takaafo. En hawaii, partiellement en maori, à Tahiti et à Nukahiva, *h* représente un ancien **f*. *f* est représenté à Fakaafo, Samoa, Tonga, Paumotu, Tahiti et aux Marquises ; *l*, à Fakaafo, Samoa, Tonga et Hawaii, et *r*, dans les autres dialectes, celui-ci à l'exclusion de celui-là et réciproquement ; aucun dialecte ne possède les deux liquides. *r* polynésien est nettement dental au point qu'on croit quelquefois entendre un *d* faible. *v* est représenté dans tous les dialectes à l'exception du maori et du hawaii qui ont *w*. *ṅ* est éga-

lement représenté dans tous les dialectes à l'exception de Tahiti où il est passé à l'occlusive glottale, de Hawaii où il est passé à *n* et de Nukahiva où il est représenté par *n* ou *k*. *n* et *m* se sont maintenus dans tous les dialectes du groupe. Les sonores *g*, *d*, *b* et les prépalatales *tš* et *dž* ne sont attestées qu'en tonga-tabu ; on les croit empruntées à la langue de l'île de Fidji voisine, car elles n'apparaissent nulle part ailleurs en polynésien.

Le polynésien n'admet ni groupe consonantique, ni syllabe finale fermée, ni diphtongue.

Les langues modernes du groupe ont de nombreux homophones, par suite de leur tendance à éliminer les consonnes. Aux Marquises, par exemple, *ua* signifie « pluie, deux, homard, grotte, vomir, se chauffer » et marque le passé. La phonétique comparée nous permet de retrouver avec certitude l'origine de ces sept sens différents. *Ua* « pluie » < *uha*, *usa* ; *ua* « deux » < *rua* ; *ua* « homard » < *uka* ; *ua* « grotte » < *rua* ; *ua* « vomir » < **rua* (comparer maori *rua-ki* avec le même sens) ; *ua* « se chauffer » < *ura* ; *ua* « marque du passé » < *kua*.

Morphologie.

D'une langue à l'autre, tel thème radical est tantôt en fonction verbale, tantôt en fonction nominale : samoa *aña*, maori *haña* « travailler » = nukahiva *hana*, *haka* « travail ». Comme en mélanésien (voir ci-dessus, p. 443), les deux fonctions sont remplies par le même mot : maori *korero*, tahiti *orero*, hawaii *olelo* signifient en même temps « parler » et « langage, discours ».

La réduplication partielle ou intégrale du thème radical donne à la forme redoublée un sens : fréquentatif (rarotonga *kati* « mordre », *katikati* « mordre souvent »), intensif (samoa *tala* « parler », *talatala* « jaser, crier »), de simultanéité (tonga *nofo* « habiter », *nonofo* « habiter avec quelqu'un »).

Verbe. La catégorie verbale polynésienne comprend :

1° des verbes transitifs ou intransitifs, passifs, causatifs, désidératifs et réciproques.

La conjugaison des premiers verbes s'exprime ainsi pour chaque temps et mode : une particule de temps ou de mode + pronom sujet + verbe ou particule de temps + verbe + pronom sujet : Tonga *oku au alu* ou *oku alu au* « je vais » = *oku*, particule du présent : *au* « je » ; *alu* « aller » ; etc.

Le passif se forme en ajoutant au thème verbal les suffixes suivants : samoa *-a*, *-ia*, *-iña*, *-ñia*, *-tia*, *-fia* ; tonga *-i*, *-ia*, *-na*, *-kia*, *-tia*, *-hia* ; rarotonga *-a*, *-ia*, *-kia*, *-mia* ; maori *-a*, *-ia*, *-na*, *-ina*, *-ña*, *-kia*, *-kina*, *-tia*, *-hia*, *-mia*, *-ria* ; tahiti *hia* ; hawaii *-a*, *-ia*, *-hia*, *-lia* ; Marquises *-a*, *-ia*, *-tia*, *-hia*.

Le causatif s'exprime par la préfixation au thème verbal des préfixes : tonga *faka-*, samoa *fa'a-*, maori *whaka-*, tahiti *fa'a*, *ha'a* ; hawaii *ho'o-* ; Marquises *faka-*, *haka-*, *fa'a-*, *ha'a-* ; rarotonga *aka*.

Le désidératif s'exprime par des préfixes : samoa *fia-*, tonga *fia-*, *fie-* ; maori *hia-* : *inu-* « boire », samoa *fia-inu-*, maori *hia-inu* « désirer boire ».

Les verbes de réciprocité et de simultanéité se forment en ajoutant au thème verbal le préfixe *fe-* et, dans nombre de cas, le préfixe précédent et les suffixes *-aki*, *-faki*, *-laki*, *-taki* ou *-ñaki* : tonga *tio* « voir, regarder », *fe-tio-faki* « se regarder l'un l'autre ».

2° les verbes de proximité ou d'éloignement qui sont caractérisés par la postposition au verbe des particules *nei* indiquant la proximité ou *na* (maori et rarotonga *ana*, hawaii *la*), l'éloignement.

3° les verbes indiquant la direction de l'action à l'aide de particules préfixées ou suffixées. Ces particules marquent si l'action est dirigée en haut, en bas ou sur le côté, et en même temps si elle a pour objet la première personne (moi, nous), la seconde (toi, vous), ou la troisième (lui, eux).

4° les verbes d'affirmation absolue qui s'expriment à l'aide de la particule *kua*, *ua* dans toutes les langues du groupe (sauf maori *ka*) ; d'affirmation conditionnelle avec les particules *ra*, *la* qui se postposent au verbe. Les verbes négatifs sont formés par les particules négatives : samoa *le*, tonga *ikai*, maori *te*, *kore* (au présent

et au futur), *kihai* (au prétérit) ; hawaii *aole* ; nukahiva *akoe*, *a'oe*, *koe*, *auma* ; tahiti *eere*, *eete*, *eore* (au présent), *eima*, *eina*, *eita*, (au futur), *aore*, *aima*, *aiña*, *aipa*, *aita* (au prétérit).

D'autres particules marquent les temps et modes du verbe.

ARTICLE. L'article polynésien possède une forme plurale que ne connaissent pas les autres groupes de la famille malayo-polynésienne :

	Singulier	Pluriel
Maori	*te*	*ña*
Tonga	*ha*	*ñahi*
Hawaii	*ke*	*na*
Tahiti	*te*	*te mau*
Marquises...	*te*	*te tau*

NOM. Comme dans un grand nombre de langues du domaine, le complément du nom devait se marquer par la postposition du complément au nom qui le régit. Le groupe conserve encore un certain nombre de formations du type : Maori *tuke mata*, littéralement arc[de l']œil = sourcil. Mais le polynésien a évolué et use plutôt, dans la langue moderne, d'une préposition pour marquer le génitif, en employant en même temps l'article tant avec le nom qu'avec son complément. Il en est de même pour les autres cas, et on a pu constituer ainsi la sorte de déclinaison suivante de maori *te tañata* « l'homme » :

	Singulier	Pluriel
Nominatif....	*te tañata*	*ña tañata*
Génitif.......	*o tañata*	*o ña tañata*
	a »	*a* »
Datif.........	*ki* »	*ki* »
Accusatif......	*i* »	*i* »
Instrumental...	*i* »	*i* »
Ablatif.......	*e* »	*e* »

Les mêmes faits s'observent dans les autres langues du groupe avec des articles et des prépositions différents.

Pronoms. Les pronoms personnels s'expriment aux trois nombres : singulier, pluriel et duel, avec une première personne du pluriel inclusive et exclusive tant au pluriel qu'au duel, par le thème pronominal : *aku, koe, ia* ou *ua* au singulier ; *ta-toru, ko-toru, ra-toru* au pluriel, et *ta-rua, ko-rua, ra-rua* au duel (*rua* = « deux » et *toru* = « trois »).

Le possessif est indiqué par un complexe signifiant « celui de moi, de toi, de lui, etc. » qui se postpose ou se prépose au nom possédé.

Noms de nombre. De 1 à 10, le polynésien reproduit, avec les alternances phonétiques attendues, la numération décimale de l'indonésien :

	Fakaafo	Samoa	Tonga	Maori	Rarotonga	Tahiti	Hawaii	Marquises
1	*tasi*	*tasi*	*tahu*	*tahi*	*tai*	*tahi*	*kahi*	*tahi*
2	*lua*	*lua*	*na*	*rua*	*rua*	*rua*	*lua*	*ua*
3	*tolu*	*tolu*	*tolu*	*toru*	*toru*	*toru*	*kolu*	*tou*
4	*fa*	*fa*	*fa*	*wa*	*a*	*ha*	*ha*	*fa, ha*
5	*lima*	*lima*	*nima*	*rima*	*rima*	*rima*	*lima*	*ima*
6	*ono*	*ono*	*ono*	*ono*	*ono*	*ono*	*ono*	*ono*
7	*fitu*	*fitu*	*fitu*	*witu*	*itu*	*hitu*	*hiku*	*fitu, hitu*
8	*valu*	*valu*	*valu*	*waru*	*varu*	*varu*	*walu*	*vau*
9	*iva*	*iva*	*hiva*	*iwa*	*iva*	*iva*	*iwa*	*iva*
10	*fulu*	*sefulu*	*hoñofulu*	*ñahuru*	*ñauru*	*ahuru*	*'umi*	*onohuu*

BIBLIOGRAPHIE

Friedrich Müller. *Grundriss der Sprachwissenschaft*. Vienne, 1882, in-8°, t. II, 2e part.

Edward Tregear. *The Maori-polynesian comparative dictionary*. Wellington (Nouvelle-Zélande), in-8°, s. d., aux pages x-xi, on trouvera une bibliographie du polynésien.

LES LANGUES PAPOUES

Entre les îles occidentales à langues mélanésiennes et micronésiennes et les dernières îles orientales à langues indonésiennes, s'étend un domaine insulaire qui comprend la Nouvelle-Guinée et quelques îles voisines.

Au Nord-Ouest de la Nouvelle-Guinée, les langues du groupe insulaire de Halmahera et des îles voisines sont trop imparfaitement connues pour être inscrites dans l'indonésien ou telle autre famille linguistique. Les langues de la Nouvelle-Calédonie qu'on dénomme généralement « papoues », d'après les indigènes à peau noire dits Papous qui l'habitent, sont également mal connues et ne peuvent entrer dans une classification générale. Des constatations faites en différentes régions, on peut seulement conclure qu'il n'y a pas unité linguistique en Nouvelle-Guinée. « A en juger d'après la langue mafor, dit Friedrich Müller dans son *Grundriss der Sprachwissenschaft* (t. I, 2e part., p. 30), les langues papoues sont *essentiellement différentes* (*grundverschieden*, souligné dans le texte) des langues malayo-polynésiennes. Il leur manque notamment la concordance phonétique des pronoms possessifs *suffixes* qui fait le mieux et le plus nettement reconnaître que les langues mélanésiennes sont apparentées aux langues malayo-polynésiennes quoiqu'elles s'écartent du type initial ». Après avoir cité cette opinion de F. Müller, H. Kern reprit la question dans un mémoire intitulé : *Over de verhouding van het Mafoorsch tot de Maleisch-polynesische talen* (*Sur la relation du Mafor avec les langues malayo-polynésiennes*, dans *Actes du 6e congrès international des Orientalistes tenu en 1883 à Leide* (section polynésienne, Leide, 1885, in-8°, p. 217-272), et conclut ainsi : « Dans les pages précédentes

(p. 220-271), un tiers environ du vocabulaire mafor qui nous est connu par les travaux des missionnaires, a été discuté. Après en avoir retranché un certain nombre de mots qui, à première vue, se reconnaissent comme des emprunts au malais, etc., et d'autres mots qui sont vraisemblablement d'origine étrangère, reste un nombre énorme de thèmes radicaux et de mots dérivés qui sont identiques ou apparentés à d'authentiques mots malayo-polynésiens. Parmi ces derniers, figurent les mots les plus simples tels que les termes pour « manger, boire, dormir, poisson, oiseau, feu », les noms de nombre, les pronoms. De plus, il est apparu que le mafor a en commun avec les langues malayo-polynésiennes les préfixes les plus usités : *a*, *aṅ*, *ma*, *ba* ou *bĕr*, *maṅ*, *pa* (*fa*) ; les infixes *in*, *um*, *m* ; les suffixes *an*, *ĕn*, *akĕn* (*épen*), *i*. Comme presque toute la grammaire des langues malayo-polynésiennes est contenue dans l'enseignement des préfixes, infixes et suffixes, on peut ainsi dire que le mafor, autant qu'on peut le prendre actuellement pour représentatif du groupe, et le groupe linguistique papou en général, avaient initialement les mêmes formes grammaticales et les possèdent encore en partie, que la famille malayo-polynésienne. Par là, ajoute Kern, l'exacte relation du papou avec le malayo-polynésien n'est pas déterminée avec précision ; cependant, autant qu'on peut l'affirmer avec certitude, le mafor n'est pas plus éloigné du malayo-polynésien que, par exemple, l'anglais de l'hindi ou le suédois du persan moderne. La parenté généalogique de la langue des Papous avec celle des Malayo-polynésiens est, pour cette raison, indéniable. On peut maintenant soutenir que les deux races sont dissemblables et, par conséquent, en déclarant que leurs langues sont apparentées, avoir recours à l'hypothèse que les Papous ont initialement appris à parler des Malayo-polynésiens, ou à une autre conjecture ; mais le fait de la parenté linguistique n'est pas, à mon avis, discutable. »

En 1908, le Dr N. Adriani a publié dans la *De Zuidwest Nieuw-Guinea-expeditie van het Kon. Ned. Aardrijkskundig Genootschap, 1904-1905* (L'expédition dans le Sud-Ouest de la Nouvelle-Guinée de la Société royale néerlandaise de géographie), un article intitulé : *Eenige opmerkingen over de Mĕraukĕ-taal naar aanlei-*

ding der woordenlijst van Contr. J. Seyne Kok (Quelques observations sur la langue de Mĕraukĕ [dans le Sud-Est de la Nouvelle-Guinée hollandaise] à l'occasion du vocabulaire [recueilli] par le Contrôleur J. Seyne Kok, Leyde, 1908, p. 637-666). Après une étude du vocabulaire en question suivent des indications sur : alphabet, accent tonique ; formation des mots par réduplication, par préfixes et suffixes ; substantifs indépendants, prépositions, verbes, noms de nombre (les indigènes de Mĕraukĕ n'ont que 1 et 2 ; 3 est 2 + 1, 4 = 2 + 2, 5 = 2 + 2 + 1 ; au-dessus de 5, tout nombre est indiqué par « beaucoup » ; de 6 à 10, on se sert des noms de nombre malais), pronoms, adverbes. Et le savant linguiste qu'est le Dr Adriani conclut : « Aussi bien d'après le vocabulaire que d'après les éléments grammaticaux, il est évident que la langue de Mĕraukĕ n'est en aucune façon malayo-polynésienne (p. 661) ».

D'après les *Reports of the Cambridge anthropological expedition*, t. III, *Linguistics* par Sidney H. Ray et les précédents travaux du même auteur (*A study of the languages of Torres Straits*, Dublin, 1893 ; *A comparative vocabulary of the dialects of British New-Guinea*, Londres, 1895), on désigne sous le nom de langues papoues les langues qui n'appartiennent ni au groupe australien, ni au groupe mélanésien, c'est-à-dire en ce qui concerne celui-ci, à la famille malayo-polynésienne. Comme nous ignorons encore ce que sont exactement les langues australiennes, cette classification provisoire est sans valeur. Les langues dites papoues, avec le sens imprécis qu'a donné M. Sidney Ray à ce terme, ne sont pas restreintes à la Nouvelle-Guinée. Deux missionnaires en ont découvert d'autres dans une île de la Mélanésie : l'île Bougainville de l'archipel des Salomon. Le Père J. Grisward a publié dans l'*Anthropos* des *Notes grammaticales sur la langue des Teleï, Bougainville, Iles Solomones* (sic) (t. V, 1910, p. 82-84 et 381-406), et, dans la même revue, le Père J. P. Rausch a étudié *Die Sprache von Südost-Bougainville, Deutsche Salomon Inseln* (t. VII, 1912, p. 105-134, 586-616, 964-994). Les langues en question sont parlées sur les côtes méridionales de l'île Bougainville par les tribus maritimes suivantes : Narioi, Koromira, Koianu, Telei,

Motuna, Kongara, Evo. Entre ces indigènes parlant une langue papoue, sont intercalés des Torau et des Banuni dont la langue est mélanésienne. Ces langues papoues possèdent l'*m* spécial qui a été précédemment noté *ṃ* (voir ci-dessus, p. 438). Le Père Grisward n'en signale pas l'existence en Telei.

En résumé, en Nouvelle-Guinée coexistent des langues nettement malayo-polynésiennes et d'autres dites papoues qui n'ont rien de commun avec les précédentes ni avec les langues australiennes voisines. A Bougainville, en Mélanésie, les unes et les autres voisinent également. Les langues dites papoues comme les langues australiennes sont trop imparfaitement connues pour être classées. On ne peut qu'en signaler l'existence.

BIBLIOGRAPHIE

Aux ouvrages précités, ajouter :

Wissenschaftliche Ergebnisse einer amtlichen Forchungsreise nach dem Bismarck-Archipel in Jahre 1908 ; t. II, *Beiträge zur Völker- und Sprachenkunde von Deutsch-Neuguinea*, par Georg FREDERICI, Berlin, 1912, in-4° ; t. III, *Untersuchungen über eine melanesische Wanderstrasse*, *ibid.*, 1913. Hugo ZELLER, *Deutsch-Neuguined und meine Ersteigung des Finisterre-Gebirges*. Stuttgart, Berlin, Leipzig, 1891, in-8° (notamment chap. XVI, p. 352 et suiv., et un vocabulaire de 300 mots en 29 langues papoues de la Nouvelle-Guinée et de l'archipel Bismarck allemands et 16 de la Nouvelle-Guinée anglaise).

Gabriel FERRAND.

LANGUES DE L'AUSTRALIE

Comme il n'y a pas de linguiste français qui ait étudié ce domaine, on se bornera ici à résumer brièvement les travaux du P. W. Schmidt qui vient de publier des exposés d'ensemble sur la question :

Die Gliederung der Australischen Sprachen, Vienne, 1919 (réunion d'une série de mémoires parus dans *Anthropos*, de 1912 à 1919) [avec bibliographie instructive].

Die Personalpronomina in den Australischen Sprachen (Denkschriften de l'Académie de Vienne, Phil.-hist. Kl. LXIV, 1).

Au surplus, la linguistique australienne n'est qu'à ses débuts. Une petite part seulement des parlers de l'Australie est décrite, la plupart du temps d'une manière tout à fait sommaire, et le P. W. Schmidt a dû fonder son étude presque uniquement sur des comparaisons de vocabulaire.

Ce n'est qu'un premier défrichement du sujet. Il résulte de là que, dans l'état actuel des connaissances, la question d'une parenté des parlers de l'Australie avec ceux d'autres régions ne saurait être abordée.

Réparties sur toute la surface d'un continent où la population est en général peu dense, les populations indigènes sont tout à fait clairsemées et en voie de régression ou même de disparition ; on ne compte guère plus qu'une trentaine de milliers d'indigènes, alors que la population totale s'élève au chiffre de près de cinq millions. Bien des parlers sont appelés à disparaître avant d'avoir été décrits même d'une manière sommaire. Car un

petit vocabulaire, tel que ceux auxquels on est souvent réduit, ne peut passer même pour un commencement de description.

Ces parlers n'ont même pas de nom fixe. Les uns sont nommés d'après le nom de la tribu qui les emploie, d'autres d'après la région géographique où habitent les indigènes, d'autres d'après la forme de tel ou tel mot du parler, de sorte qu'il serait vain d'énumérer les noms cités par le P. Schmidt. Pour plus ample étude il faut se reporter à cette liste complète, avec les indications bibliographiques jointes à chaque nom. Ici la carte seule a été reproduite (planche 13). Elle suffit à donner une idée des noms adoptés.

Quant aux classements proposés, et surtout quant aux hypothèses énoncées sur l'histoire de ces parlers, les bases en sont si étroites et si fragiles qu'il serait imprudent de les reproduire. Du reste, les faits sont complexes. Ainsi, tandis que les vocabulaires des parlers de la région du cap York et de la région Centrale diffèrent beaucoup de ceux de l'Australie du Sud, les pronoms personnels concordent en une large mesure. On entrevoit donc une certaine unité originelle des parlers de l'Australie. En l'état présent des choses, les parlers du Sud offrent des concordances notables, tandis que les parlers du Centre et du Nord divergent entre eux et d'avec ceux du Sud.

Il importe beaucoup de faire sur ces parlers une enquête approfondie avant qu'ils ne disparaissent.

A. Meillet.

LES LANGUES
DU
SOUDAN ET DE LA GUINÉE[1]

Généralités

a) *La famille négro-africaine.*

Les langues parlées par les populations noires de l'Afrique ont été l'objet de bien des essais divers de classification. Celui qui paraît avoir eu, en ces dèrniers temps, les préférences générales consiste à mettre à part les langues dites « bantou » et à grouper les autres sous l'étiquette de « langues soudanaises ».

En réalité, cette distinction entre langues « bantou » et langues « soudanaises », ou, si l'on préfère, entre les langues nègres qui sont cataloguées comme bantou et celles qui ne le sont pas, n'est pas absolue. Les unes et les autres, pour autant qu'il est possible de se prononcer sur des parlers dont on ignore l'origine et l'évolution historique, semblent bien appartenir à une même famille linguistique, ainsi qu'il sera facile de s'en rendre compte en parcourant les pages qui vont suivre et en se reportant au chapitre consacré d'autre part aux langues bantou.

Par les points les plus caractéristiques de leur morphologie et de leur syntaxe, par l'ensemble des lois phonétiques qui les régissent, par les éléments formatifs de leur vocabulaire, les unes

1. Les noms de langages ou de populations imprimés en **égyptiennes** ou en romaines doivent se lire selon les lois ordinaires de la prononciation française, ainsi que les noms portés sur la carte. Les mêmes noms ou les mots indigènes, imprimés en *italiques*, se liront selon les lois de la transcription phonétique adoptée pour l'ensemble du présent ouvrage ; on notera de plus que *ḛ*, *ḭ* et *ṵ* représentent un *e*, un *i* et un *u* nasalisés, distincts des voyelles nasales ordinaires telles que *ã*, *ẽ*, *õ*, et que b et d, figurés en romaines dans le corps d'un mot en *italiques*, représentent des consonnes spéciales, dites « claquantes », sur la nature exacte desquelles on n'est pas encore définitivement fixé.

et les autres présentent des affinités telles qu'il paraît difficile de les dissocier. Bantou ou non bantou, au Soudan comme sur les côtes de Guinée et dans les régions subéquatoriales, toutes les langues parlées par les populations d'Afrique qui appartiennent sans conteste à la race noire, ainsi que par quelques populations plus ou moins négroïdes sur l'origine desquelles on est moins nettement fixé, constituent un ensemble dont l'unité devient de plus en plus apparente à mesure que l'étude en est poussée plus avant et auquel on pourrait donner, provisoirement tout au moins, le nom de « famille négro-africaine ».

Ce qui autorise à traiter à part les langues bantou, c'est que ces langues forment un groupe nettement défini de la famille négro-africaine. Alors que ce groupe embrasse près de la moitié de l'Afrique noire et présente, depuis l'équateur jusqu'au Cap de Bonne Espérance et des rives du Congo à celles du Zambèze — exception faite de l'enclave formée par les parlers des Hottentots et des Bushmen — un aspect bien compact, on se trouve, dans les parties centrale et occidentale du continent africain, en face d'un tel chaos qu'il semble impossible de n'y pas compter moins de seize groupes différents, pour retrouver dans chacun une homogénéité à peu près comparable à celle que présentent entre elles les nombreuses langues bantou. Encore le tout formé par chacun de ces seize groupes, dont certains n'occupent qu'une superficie extrêmement restreinte, est-il loin, dans la plupart des cas, de se présenter sous un aspect aussi manifestement unique que le tout formé par le groupe bantou.

Laissant de côté ce dernier (étudié dans un autre chapitre), je parlerai seulement ici des seize autres groupes de la famille négro-africaine, c'est-à-dire des langues parlées, d'une façon générale, au Soudan et dans la Guinée septentrionale et moyenne (par 45 à 50 millions d'individus).

b) *Difficultés et caractère provisoire de la classification proposée.*

Il est malaisé de faire, même à grands traits, la grammaire comparée de ces langues et surtout d'en établir une classification généalogique.

D'abord notre documentation est presque uniquement contemporaine.

Les fameuses inscriptions de Philœ et de Méroé, pour l'instant, ne nous servent de rien, puisqu'on n'a pas encore pu déterminer si la langue dans laquelle elles ont été rédigées était celle des Blemyes, auxquels on rattache, assez hypothétiquement d'ailleurs, les Bedja de nos jours, ou bien celle des Nobades, ancêtres supposés des Nouba ou soi-disant tels. En d'autres termes, nous ignorons si le texte de ces inscriptions appartient à une langue chamito-sémitique ou à une langue négro-africaine [1].

L'unique mot rapporté par le carthaginois Hannon de son voyage à la côte occidentale d'Afrique ne nous est connu que par l'incertaine transcription dont nous avons tiré le nom du « gorille » ; on en pourrait seulement rapprocher la racine *gor*, *kor* ou *gur*, signifiant « homme » dans plusieurs langues actuelles du bas Sénégal.

Il faut descendre jusqu'à l'époque du IVe au VIIe siècle après Jésus-Christ pour rencontrer des documents de quelque importance : je veux parler des ouvrages religieux rédigés en langue nubienne et écrits en caractères coptes. La valeur linguistique en est médiocre. En effet, outre que la graphie employée ne convient nullement à l'usage qui en a été fait, ces textes ont été composés par des étrangers ou tout au moins par des hommes de culture étrangère, par des savants ; ils sont la traduction d'écrits étrangers relatifs à des choses pour l'expression desquelles le nubien de l'époque devait manquer de vocables appropriés. La langue en est savante, artificielle ; barbarismes et solécismes n'y manquent peut-être pas.

Au XIe siècle, on trouve quatre mots soudanais et une expression cités par El Bekri ; puis, au XIVe, Ibn Khaldoun nous donne cinq mots et Ibn Batouta seize, plus une proposition nominale,

1. A titre d'indication sur la langue des inscriptions méroïtiques et l'origine de leur écriture, voir : H. SCHUCHART, *Das Meroitische* (*Wiener Zeitschrif für die Kunde des Morgenlandes*, XXVII, 1913) ; A. H. SAYCE, *The origin of the Meroitic alphabet* (*Proceedings of the Soc. of Biblical Archæology*, 1914, pp. 117-180) ; R. EISLER, *Zu Demokrits Wanderjahren* (*Archiv für Philosophie*, XXIV, 1918, p. 204), et ci-dessous p. 477.

recueillis par lui sur place. J'ai étudié ailleurs ces quelques matériaux et, mettant à part trois vocables que je n'ai pu identifier nettement, j'y ai reconnu, tels qu'ils se disent aujourd'hui dans les mêmes régions, deux mots peuls, quatre mots sarakollé, deux mots et une expression songoï, quatorze mots et une proposition mandingues. Donc, dès la fin de notre moyen âge, l'état linguistique de cette partie du Soudan ressemblait à l'état actuel.

A partir du XVIIe siècle, la documentation devient plus fournie. Des vocabulaires sont recueillis au Sénégal et à la Côte d'Or, des mots et des phrases sont cités par les voyageurs. Tout cela cependant se réduit à peu de chose, présente peu de garanties, et, mis à part quelques très rares travaux d'une réelle valeur, il faut arriver vers le milieu du XIXe siècle pour entrer dans la période de recherches sérieuses à laquelle a succédé enfin, il y a seulement une trentaine d'années, une période de mise au point qui n'est encore qu'à ses débuts.

En somme, nous ne disposons que d'une documentation contemporaine. Est-elle au moins suffisamment sûre et abondante ? Ni l'un ni l'autre.

De quoi se compose-t-elle en effet ? Pour le plus grand nombre des langues, de simples listes de mots, recueillis souvent sans méthode auprès d'informateurs fréquemment douteux, par le canal d'un double ou d'un triple interprète, et transcrits selon les caprices d'une oreille plus ou moins inexpérimentée. Je ne conteste pas le mérite de ceux qui nous ont rapporté ces vocabulaires, sans lesquels nous n'aurions rien, mais quelle base fragile et décevante ils constituent pour des études linguistiques !

Nous avons, il est vrai, de quelques langues, des grammaires dont certaines sont bien faites. Mais la plupart sont des accommodations de la grammaire de Lhomond, c'est-à-dire qu'on s'y est évertué à faire entrer une langue négro-africaine dans le cadre d'une langue indo-européenne. Et, malheureusement, ce défaut capital n'est que bien rarement racheté par de copieux exemples ou des textes suivis. Exemples et textes présentent d'ailleurs les mêmes caractères d'incertitude que les vocabulaires : le noir ignorant et illettré par lequel on se fait dicter un récit trouve fasti-

dieuse cette besogne inaccoutumée dont il ne saisit pas la portée ; il cherche à l'abréger et, dans ce but, évite avec soin les locutions qu'il suppose ignorées de son collaborateur européen, afin de n'avoir pas à les lui expliquer ; il tient à n'employer que des tournures connues et toujours les mêmes ou, ce qui est pis, il émet des phrases, non point conformes au génie de sa langue, telle qu'il la parle avec ses compatriotes, mais coulées dans le moule d'une espèce de sabir à l'usage des relations entre Européens et indigènes.

Les bons textes, voilà ce qui nous manque le plus et ce qui, pourtant, est le plus indispensable au linguiste. Nous en possédons d'excellents et abondants en haoussa et, depuis une époque toute récente, en peul, parce que ces deux langues sont écrites, au moins occasionnellement, par plusieurs des gens qui les parlent, ce qui a permis de faire rédiger des contes ou des récits au lieu de se les faire dicter. Mais si l'on songe que, sur les 435 langues environ qui ont cours au Soudan et en Guinée, il n'y en a que cinq ou six qui s'écrivent, on conviendra que grande est la difficulté de se procurer de bons textes en quantité suffisante. Il existe bien, en assez grand nombre, des traductions de la Bible et du catéchisme, mais, sauf exceptions, cette catégorie de documents a peu de valeur au point de vue linguistique : la plupart de ces traductions ont les défauts des textes nubiens du IVe au VIIe siècle dont je parlais plus haut ; elles nous livrent la langue des missionnaires et non pas celle des indigènes.

Et que dire des parlers dont nous ne savons que le nom ? Il en est aussi dont nous ne soupçonnons pas l'existence.

Notre connaissance des langues négro-africaines est tellement imparfaite que toute étude comparée de ces langues, faite aujourd'hui, ne peut prétendre à autre chose qu'à poser des jalons pour l'avenir. C'est dire que la classification à laquelle je suis arrivé, après trente ans de pratique, d'étude et de tâtonnements, n'est qu'une classification incomplète et provisoire, destinée à guider les recherches plutôt qu'à résoudre le problème.

c) *Caractères communs aux divers groupes de langues négro-africaines.*

1° *Classes nominales.* — Le système des classes nominales, tel qu'il apparaît dans son intégrité, repose sur une répartition des êtres et objets et, postérieurement sans doute, des abstractions en un certain nombre de catégories, d'après une classification dont les règles tantôt se manifestent clairement à notre esprit et tantôt nous échappent, mais semblent répondre chez les noirs à des conceptions propres. En général, les êtres humains constituent l'une de ces catégories ; l'eau, les liquides et les corps facilement fusibles ou liquéfiables en forment une autre ; une troisième comprendra le bois et les végétaux ligneux, une quatrième l'herbe et les végétaux herbacés ou fibreux, une cinquième la terre avec ses divers aspects et tout ce qui provient d'elle, une sixième ce qui se rapporte à telle ou telle saison de l'année, etc. Tous les noms servant à désigner des êtres ou objets de l'une de ces catégories constituent ensemble une classe grammaticalement distincte, caractérisée par une syllabe spéciale qui lui est propre, qui sert de pronom à tous les noms de la classe, qui est employée comme déterminatif de ces mêmes noms et qui, soit sous sa forme pronominale soit en subissant telle ou telle modification peu profonde, se préfixe ou se suffixe au radical de chacun de ces noms et à celui de l'adjectif qui le qualifie.

Ce système a dû, à une certaine époque, fonctionner intégralement dans toutes les langues négro-africaines, tel qu'il fonctionne encore de nos jours dans un grand nombre d'entre elles, bantou et non bantou. Mais, de même que les langues indo-européennes offrent une tendance générale à se débarrasser des flexions casuelles, les langues négro-africaines ont eu et ont encore une tendance générale à se débarrasser des classes nominales. Toutes ont évolué dans cette voie, mais les unes ont évolué plus vite et plus radicalement que les autres. Aussi pouvons-nous, de nos jours, observer, dans un même groupe et surtout dans des groupes divers, à peu près tous les stades de cette évolution.

Certaines langues possèdent le système à l'état complet. D'autres, tout en conservant les affixes de classe comme déterminatifs, comme étiquettes des noms et comme symboles de l'accord de l'adjectif avec le nom, les ont remplacés, comme pronoms personnels, par une forme unique commune à toutes les classes. D'autres, allant plus loin, ont laissé tomber l'affixe anciennement accolé au nom, tout en le conservant à l'adjectif et en le gardant comme déterminatif. D'autres enfin ont rejeté même ces derniers procédés et ne présentent plus que des vestiges de l'ancien état de choses, sous la forme de quelques affixes de classe demeurés accolés à certains noms et faisant désormais, en quelque sorte, partie du radical.

Quoi qu'il en soit, il n'est sans doute pas une seule langue négro-africaine dans laquelle on ne retrouve point, à un degré plus ou moins apparent, sous une forme ou sous une autre, des traces de classes nominales. Dans quelques-unes, il est vrai, ces traces sont si ténues et si clairsemées qu'on ne les aperçoit pas au premier abord, mais un examen approfondi ne tarde pas à les mettre en lumière. Par ailleurs, dans un très grand nombre de langues dites « soudanaises », notamment dans celles du groupe kordofanien et dans plusieurs idiomes des groupes éburnéo-dahoméen, nigéro camerounien, sénégalo-guinéen, nilo-équatorien, voltaïque, en particulier dans plusieurs parlers quelquefois qualifiés bien à tort de chamitiques, tels que le peul et le massaï, les classes nominales fonctionnent aussi complètement que dans les langues bantou dans lesquelles le système est le plus visible.

Le plus souvent, la syllabe caractéristique de chaque classe se préfixe au radical, au Soudan et en Guinée comme dans le domaine bantou. Quelquefois, elle est à la fois préfixée et suffixée, ainsi qu'il arrive dans certaines langues du groupe voltaïque et du groupe sénégalo-guinéen et dans le groupe bas-nigérien. Enfin, dans la plupart des parlers du premier de ces groupes et dans l'un au moins de ceux du second (le peul), elle est suffixée au radical. Ces modalités ne me semblent pas constituer un élément important de différenciation, surtout étant donné que la syllabe suffixée dans une langue correspond sou-

vent de façon très nette au préfixe de même classe dans une autre langue.

2° *Manque de distinction essentielle entre le nom et le verbe dès que le système des classes nominales cesse de fonctionner à l'état complet.* — Tant que chaque nom possède son préfixe ou son suffixe de classe et que la notion du rôle joué par ce préfixe ou ce suffixe ne s'est point perdue, le nom se trouve, par là même, nettement distingué du verbe. Mais c'est là le seul élément essentiel de différenciation entre le nom et le verbe dans les langues négro-africaines. Dès qu'il disparaît, c'est-à-dire dès que l'affixe de classe est tombé ou qu'il s'est incorporé au radical, il n'y a plus rien qui permette de reconnaître que tel mot est un nom ou un verbe, en dehors de la place qu'il occupe dans la phrase et en dehors aussi, bien entendu, des cas où le verbe s'accompagne d'un affixe de conjugaison ou de celui où l'on a affaire à un nom dérivé qui ne saurait, de par la nature de son affixe de dérivation, être pris pour un verbe. Comme un radical verbal peut s'employer seul, ne serait-ce qu'à l'impératif, et comme un radical nominal peut être par lui-même un mot dans les langues qui ont rejeté les affixes de classe, il s'ensuit que l'évolution dans ce sens marche de pair avec l'évolution des classes nominales vers la disparition ; dans les langues où ces dernières ont totalement disparu en fait, comme en mandingue, on a parfois recours dialectalement à des artifices pour distinguer le nom du verbe, mais, en réalité, rien ne les distingue plus essentiellement sauf leur place respective, et la syntaxe est obligée de venir au secours d'une morphologie absente ou inopérante.

3° *Aspects verbaux.* — Les conceptions des noirs africains ne correspondent pas généralement aux nôtres et il n'est pas possible, sans égarer les idées, de donner les noms de nos temps ou de nos modes aux divers aspects du verbe dans les langues négro-africaines. D'autre part, ces conceptions spéciales se retrouvent identiquement les mêmes dans toutes ces langues, bien qu'elles y soient rendues de façons variées.

Dans toutes, en mettant à part l'infinitif, qui est un nom, et

les participes, qui sont des noms ou des adjectifs, le verbe comporte trois aspects essentiels, les autres dérivant le plus souvent des premiers ou étant obtenus à l'aide d'auxiliaires, ou de formes spéciales données aux pronoms sujets.

L'un de ces trois aspects a surtout, par rapport à notre façon de voir, une valeur négative; il indique avant tout que l'action n'est pas accomplie ou que l'état n'est pas acquis; il peut donc représenter soit notre présent, soit notre futur, et il représente en effet l'un et l'autre dans une certaine mesure, ainsi que le présent d'habitude et le présent de narration employé pour le passé; lorsqu'il y a lieu de le traduire par le futur, il marque plus particulièrement une réalisation prochaine ou inéluctable de l'action ou de l'état, dans le sens, par exemple, de « il va partir » ou « il partira nécessairement, il doit partir ». Je donne à cet aspect verbal, faute de mieux, le nom d'« aoriste ».

Un second aspect, que j'appelle « parfait », implique que l'action est achevée ou que l'état est acquis et dure encore. Il traduira donc, d'une manière générale, le passé de nos verbes d'action et le présent de nos verbes d'état : « il a mangé, il est parti » se rendront par le parfait dans les langues négro-africaines, et « je sais, je comprends » se rendront aussi par le parfait, attendu que ces expressions sous-entendent que « j'ai acquis » la connaissance ou l'intelligence d'une chose, de même que « il est grand, il me plaît », parce que, dans ce dernier cas, le fait d'être grand ou d'être agréable est effectivement acquis et dure encore.

Le troisième aspect se réfère à une action ou à un état qui s'accomplit ou s'acquiert, soit dans le passé, soit dans le présent, soit dans le futur, sous l'influence d'une autre action ou d'un autre état pouvant résulter de la volonté du sujet, mais pouvant aussi lui être étrangère. Je l'appelle « injonctif »; en plus de notre subjonctif, il rend aussi l'impératif, l'optatif et, en général, notre infinitif complément d'un autre verbe ou en dépendant.

L'aoriste négro-africain peut exiger, dans certains cas, d'être traduit par l'un de nos passés, car il peut exprimer une action qui, tout en étant passée en ce qui concerne le début de son accomplissement, n'est pas définitivement achevée et par conséquent

est encore en cours soit au moment où l'on parle soit au moment dont il s'agit : par exemple, si l'on dit, en parlant d'un homme encore vivant, « il a fait de grandes choses toute sa vie » ; ou en ce qui concerne le premier verbe d'une phrase comme celle-ci : « il travaillait aux champs lorsque nous l'avons rencontré ».

Inversement, le parfait sera fréquemment traduit par notre présent ou notre futur lorsqu'il s'agira d'une action antérieure à celle exprimée par la proposition principale : dans des phrases comme « viendras-tu me voir si je te donne un présent ? », « je te donnerai un présent quand tu travailleras bien », les verbes « donne » et « travailleras » se mettront au parfait, parce que j'aurai accompli l'action de donner au moment où tu viendras et que tu auras accompli celle de travailler au moment où je te récompenserai.

Dans la plupart des langues négro-africaines, il existe d'autres aspects verbaux qui expriment un passé relativement à l'un des aspects principaux, tels qu'un « imparfait » correspondant à l'aoriste, un « prétérit » correspondant au parfait : ceux-là pourraient être appelés des « temps ». On les rend le plus souvent à l'aide d'un affixe spécial venant s'ajouter à la forme de l'aspect principal. On rencontre aussi des aspects secondaires tels qu'un présent absolu, un futur ordinaire, un conditionnel, etc., qui tantôt se forment comme les aspects principaux, tantôt dérivent de ceux-ci et tantôt se construisent au moyen d'un auxiliaire ou d'un pronom sujet spécial qui, alors, est indispensable et subsiste à la 3e personne, même s'il y a un nom sujet.

Mais partout, l'aoriste, le parfait et l'injonctif forment la base de la conjugaison. Parfois même, ils sont les seuls aspects verbaux qui existent. On les forme en général à l'aide d'affixes de conjugaison ; souvent l'un d'entre eux ou même deux d'entre eux sont constitués par le simple radical verbal ; il en est ainsi presque partout de l'injonctif à la 2e personne du singulier, correspondant à notre impératif.

4° *Syntaxe de position.* — L'absence de toute flexion casuelle nécessite la subordination du rôle joué par les mots dans la phrase à la place respective qu'ils y occupent : aussi toutes les

langues négro-africaines ont une syntaxe de position. A vrai dire, quelques-unes font usage de particules suffixées ou préfixées qui, dans une certaine mesure, ont une valeur analogue à celle des flexions casuelles dans d'autres familles linguistiques ; toutefois ces particules ne peuvent rendre que quelques cas (possessif, génitif, locatif, datif), rarement l'accusatif, plus rarement encore le nominatif ; leur rôle, en outre, n'est qu'accessoire, ainsi que le prouve le fait qu'elles sont souvent omises dans les langues même qui en font le plus usage. Ce qui constitue essentiellement la syntaxe dans l'ensemble des langues négro-africaines, ce qui marque qu'un mot est un nom, un adjectif, un verbe, qu'un nom est sujet ou complément du verbe, ou complément d'un autre nom, c'est la place qu'occupe ce mot par rapport aux autres.

Quant à l'ordre adopté, il varie selon les groupes. C'est surtout pour la place respective du nom et de son complément, du verbe et de son complément, qu'il y a des différences contribuant précisément à faire rattacher telle ou telle langue à tel groupe plutôt qu'à tel autre.

5° *Vocabulaire*. — Les termes d'importation étrangère mis à part, les éléments formatifs du vocabulaire, racines et affixes, présentent un remarquable caractère d'unité dans l'ensemble des langues négro-africaines. Il arrive souvent que plusieurs racines servent à exprimer la même idée et que, dans ce cas, l'une de ces racines ait prévalu dans une langue tandis que l'une des autres a été adoptée de préférence par une autre langue, mais il n'est apparemment pas un seul groupe négro-africain dans lequel on ne puisse retrouver, à quelques exceptions près, tous les principaux éléments formatifs du vocabulaire d'un autre groupe quelconque, si grande que soit la distance qui sépare leurs domaines géographiques respectifs. Non seulement les racines sont similaires et proviennent manifestement d'une ancienne racine commune, mais encore les concepts exprimés par ces racines similaires sont rigoureusement identiques et les concepts qui en sont obtenus par dérivation le sont au moyen de procédés analogues donnant des résultats semblables.

Cette communauté se manifeste également dans certains pro-

noms, dans certains noms de nombre et dans les affixes des classes les mieux caractérisées (classe humaine et classe des liquides principalement).

6° *Tons musicaux.* — Certaines langues négro-africaines possèdent des tons musicaux à valeur étymologique ou à valeur grammaticale, d'autres n'en possèdent pas, et les deux cas peuvent co-exister dans le même groupe. Ce qu'il importe d'observer, c'est que, partout où les tons ont une valeur grammaticale, le ton bas marque l'affirmation, l'augmentatif, le pluriel ou la personne à qui l'on parle, tandis que le ton haut marque la négation, le diminutif, le péjoratif, le singulier ou la personne qui parle.

d) *Différences conduisant à répartir en plusieurs groupes distincts les langues négro-africaines.*

Les classes nominales, quoique apparaissant partout, ne sont pas partout au même stade, comme il a été dit plus haut. La difficulté de distinguer le nom du verbe, dès que le système des classes nominales tend à se simplifier, est générale, mais les procédés employés pour y remédier varient. La valeur des aspects verbaux est identique dans toutes les langues, mais le système de conjugaison ne l'est pas. Toutes ont recours à la syntaxe de position, mais toutes n'ont pas adopté le même ordre des mots. Les éléments formatifs du vocabulaire sont communs, mais des phonétiques différentes viennent en modifier l'aspect.

C'est par la comparaison des modalités d'application des grands principes communs que j'ai été amené à répartir les langues négro-africaines en dix-sept groupes, l'un de ces groupes étant le groupe bantou et les seize autres occupant le domaine de l'Afrique noire non bantou. A l'intérieur de chacun de ces groupes, les modalités d'application des principes communs sont les mêmes, au moins dans leurs grandes lignes, et c'est là ce qui fait l'unité du groupe.

Accessoirement, j'ai examiné dans le même esprit la façon dont se marque le pluriel des noms et la forme que revêtent les pronoms personnels.

Deux langues ayant même grammaire et même vocabulaire et appartenant incontestablement au même groupe linguistique peuvent présenter, à la même époque, des aspects phonétiques dissemblables, si elles sont parlées dans deux régions différentes, tandis qu'au contraire deux langues de systèmes grammaticaux divergents, mais parlées dans des contrées analogues, peuvent avoir la même phonétique. Aussi les caractères d'ordre purement phonétique m'ont-ils paru être ceux qui présentent en général le moins d'importance pour la classification des langues négro-africaines, surtout étant donné que nous manquons le plus souvent de certitude quant aux faits de cet ordre.

Dans chaque groupe se trouvent des langues qui font, pour ainsi dire, liaison avec quelques langues des groupes voisins. Souvent, il est difficile de tracer la démarcation, et ces espèces de soudures font apparaître plus étroite l'unité de l'ensemble. C'est ainsi que les langues les plus orientales du groupe nigéro-camerounien, telles que le tikar, le vouté, etc., et les langues nord-occidentales du groupe bantou, telles que le douala, le fang, etc., sont si voisines les unes des autres que l'on hésite à rattacher les unes ou les autres à l'un des groupes plutôt qu'à l'autre. Le même phénomène s'observe dans l'Ouest et le Sud-Ouest de l'Éthiopie entre certaines langues que je crois devoir considérer comme négro-africaines et certaines autres que l'on estime être des parlers chamito-sémitiques du groupe couchitique[1]. C'est une raison de plus pour répéter que la présente classification n'offre pas un caractère absolu ni définitif.

1. Peut-être est-ce ici qu'il y a lieu de mentionner l'existence, sur les rives de l'Ouêbi-Chabêli, de populations nègres dites *Adōn*, *Adōno*, *Šabẹ̄li* ou *Gubaḥīn*, qui vivent sous la domination des Somali de la fraction Haouiya et qui parleraient, en plus du somali, une langue négro-africaine qu'on suppose être bantou.

e) *Note sur les langues écrites.*

Dans l'ensemble, les langues du Soudan et de la Guinée sont uniquement des langues parlées et ne s'écrivent pas. Toutefois deux d'entre elles au moins possèdent un système d'écriture qui leur est propre et qui est employé sur une assez vaste échelle à l'intérieur de leur domaine : 1° le **vaï** (groupe nigéro-sénégalais) s'écrit au moyen de signes syllabiques d'invention indigène, dont l'origine remonte soit à la fin du XVIIIe siècle soit au début du XIXe ; 2° le **mom** ou bamoun (groupe nigéro-camerounien) s'écrit au moyen de signes inventés vers 1900 par Njoya, roi de Foumbân, et modifiés depuis par leur inventeur ; d'abord idéographiques, ces signes sont devenus phonétiques ; encore syllabiques aujourd'hui, ils tendent à devenir alphabétiques ; leur usage se répand parmi quelques populations voisines des Bamom ou Bamoun et parlant des langues similaires, notamment les Bagam.

Peut-être la langue des inscriptions dites « méroïtiques » était-elle le **nouba** (groupe nilo-tchadien) ou un parler dont procéderait plus ou moins directement ce dernier ; les caractères coptes ont été utilisés dans le domaine religieux, du IVe siècle au VIIe, pour la transcription du nouba ancien, et, d'après une communication récente de H. A. MacMichael, le nouba moderne s'écrirait de nos jours, de Korosko au Mahas, au moyen de signes se rapprochant à la fois de ceux de l'alphabet arabe et de ceux d'alphabets sémitiques anciens.

D'autre part, trois langues s'écrivent, au moins occasionnellement, à l'aide de l'alphabet arabe, parfois adapté au moyen de signes diacritiques supplémentaires : le **kanouri** (groupe nilo-tchadien), le **haoussa** (groupe nigéro-tchadien) et le **peul** (groupe sénégalo-guinéen). Ce procédé ne s'est pas étendu aux autres langues parlées par des populations islamisées ; les noirs musulmans lettrés, lorsqu'ils veulent écrire, le font en langue arabe ; nulle part, sauf quelques exceptions individuelles ou localisées, ils ne parlent l'arabe, mais, partout, ils ont adopté l'arabe écrit comme langue savante et langue de correspondance.

Enfin, il convient de noter que certains catéchumènes des missionnaires chrétiens écrivent leur langue maternelle au moyen de caractères romains plus ou moins adaptés aux nécessités phonétiques locales. Ce procédé toutefois n'a pas pris, en général, une grande extension.

INDICATIONS BIBLIOGRAPHIQUES SUR LES ÉCRITURES. — Voir, pour l'écriture vaï, DELAFOSSE, *Les Vaï, leur langue et leur système d'écriture* (*L'Anthropologie*, X, 1899) ; pour l'écriture mom, même auteur, *Naissance et évolution d'un système d'écriture de création contemporaine* (*Revue d'ethnographie*, 1922, 9) ; pour l'écriture nouba moderne, MACMICHAEL, *A history of the Arabs in the Sudan*, Cambridge, 1922, in-8 (2e vol., p. 328) ; pour le méroïtique, les divers travaux de F. L. GRIFFITH et notamment ses *Meroitic inscriptions*, dans *Archæological survey of Egypt*, XIX, 1911, ainsi que W. E. A. BUDGE, *The Egyptian Sûdân, its history and monuments*, 1907, 2 vol. (copieuse bibliographie), et ci-dessus p. 465 note.

Je vais maintenant examiner successivement les seize groupes négro-africains du Soudan et de la Guinée. Faute de pouvoir, dans l'état actuel de la science, adopter un ordre généalogique solide, je me suis arrêté à un ordre approximativement géographique, en allant, autant que possible, du Nord au Sud et de l'Est à l'Ouest, tant pour les groupes eux-mêmes que pour les langues dont chacun d'eux se compose.

Afin de fixer les idées, j'ai donné à chaque groupe un nom qui comporte en lui-même l'indication sommaire de la localisation territoriale du groupe. On trouvera sur la carte qui forme la planche 14 du présent volume la figuration du domaine géographique de chaque groupe, ainsi que les noms des diverses langues qui le constituent, le nom de chaque langue correspondant approximativement à l'emplacement de celle-ci [1].

1. Toutefois, ne figurent pas sur la carte les noms de dix langues sur lesquelles je n'ai pu me procurer les précisions nécessaires qu'après le tirage de cette carte. Ces langues sont : le siga (gr. nilo-tchadien) ; le kasselé et le mouâ (gr. voltaïque) ; l'ago, le likpé, l'akpafou, le bôli et le boro (gr. éburnéo-dahoméen) ; le boussa et le boussan (gr. nigéro-sénégalais).

Au cours de l'énumération des diverses langues de chaque groupe, j'ai mis entre parenthèses, pour chaque langue, une date et un nom : c'est la date à laquelle remontent les premiers renseignements obtenus et le nom de l'auteur auquel ils sont dus. L'absence de tout renseignement publié est indiquée par le mot « néant ».

INDICATIONS BIBLIOGRAPHIQUES SUR L'ENSEMBLE DES LANGUES DU SOUDAN ET DE LA GUINÉE. — En fait d'ouvrages traitant de l'ensemble des langues du Soudan et de la Guinée, ou tout au moins de plusieurs groupes de langues, on peut citer : R. N. CUST, *A sketch of the modern languages of Africa*, London, 1883, 2 vol. pet. in-8° (actuellement sans autre valeur que celle de fournir une abondante bibliographie, parfois incorrecte d'ailleurs, pour la période antérieure à 1880) ; D. WESTERMANN, *Die Sudan-sprachen*, Hamburg, 1911, gr. in-8 (essai de reconstitution d'un ancien soudanais commun par des comparaisons entre trois langues du groupe éburnéo-dahoméen, deux langues du groupe nigéro-camerounien, deux langues du groupe nilo-tchadien et une langue du groupe nilo-abyssinien) ; F. W. H. MIGEOD, *The languages of West-Africa*, London, 1911-1913, 2 vol. in-8 (étude comparative, mais trop souvent superficielle, d'un grand nombre de langues de l'Afrique occidentale) ; C. MEINHOF, *An introduction to the study of African languages*, London, 1915, in-8 (résultat de recherches sérieuses et approfondies, offrant toutefois l'inconvénient de quelques confusions entre le domaine négro-africain et le domaine chamito-sémitique) ; A. DREXEL, *Gliederung der Afrikanischen Sprachen* (en cours de publication dans *Anthropos*, où le début de ce nouvel essai de classification a paru en 1921-1922, XVI-XVII, fasc. 1-2-3).

Les seize groupes négro-africains du Soudan et de la Guinée

I. — *Groupe nilo-tchadien* (30 langues).

Ce groupe, le plus septentrional et sans doute le plus vaste, occupe la vallée du Nil depuis Assouan (1re cataracte) au Nord jusque non loin de Fachoda au Sud, s'étendant à l'Est entre le Nil et l'Atbara et au delà de cette rivière jusqu'aux limites occidentales du tigraï; à l'Ouest, il pénètre à travers le désert de Libye jusqu'au Tibesti inclus, englobant toutes les régions montagneuses et désertiques situées au Nord du Kordofan, du Darfour, du Ouadaï et du Kanem, pour s'étendre plus au Sud dans la majeure partie de ces quatre provinces et, contournant le lac Tchad au Nord, dans le Mounio et le Bornou. Cet immense domaine présente peu de solutions de continuité, mais le groupe nilo-tchadien le partage dans le bassin du Nil et à l'Est et au Nord du Tchad avec l'arabe, que parlent de nombreuses tribus, pour la plupart nomades, d'origines diverses.

Le nombre des individus ayant comme parlers maternels des langues nilo-tchadiennes ne saurait être évalué rigoureusement, mais il dépasse très certainement deux millions, qui se répartissent entre plusieurs peuples, les uns constitués par des négroïdes et des nègres (comme une fraction des Nouba, les Baria, les Kounama et l'ensemble des Zaghaoua, Anna et Toubou), les autres composés de nègres purs ou à peu près purs (les Nouba du Sud, les Kondjara, les Mâba, les Kanouri, etc.).

Certains, comme les Nouba du Nord, ont joui autrefois d'une civilisation avancée, sans qu'on puisse préciser si elle était ou non d'importation étrangère. Actuellement, on ne rencontre une civilisation relativement développée qu'au Darfour chez les Kon-

djara, au Ouadaï chez les Kodoï et les Mâba, au Bornou chez les Kanouri, c'est-à-dire dans le Sud du groupe et seulement dans celle de ses fractions qui appartient incontestablement à la race noire, tandis que les autres populations, en particulier les négroïdes du Nord, dont certains sont aussi voisins de la race blanche que de la race noire (Zaghaoua, Anna, Toubou), ne possèdent qu'une civilisation très fruste.

Les traces de classes nominales sont faibles, mais visibles. En kounama, on a un préfixe *a* pour les noms concrets, un préfixe *ko* pour les noms abstraits; en nouba, un préfixe à initiale *k* ou *g*; en baria, l'emploi du pronom de la 3ᵉ personne comme déterminatif; en kanouri et en toubou, quelques préfixes, dont un préfixe nasal pour les noms de liquides; en mâba, kodoï, mimi, mara, kondjara, rougna, etc., des préfixes assez nombreux, dont *mu* pour la classe humaine. Les pronoms de classe n'ont subsisté nulle part.

On a cherché à distinguer le nom du verbe en donnant au nom une désinence spéciale : *a* en kounama, voyelles diverses en nouba, voyelles ou consonnes diverses dans les autres langues.

Les aspects du verbe sont marqués par des suffixes; le plus souvent, une voyelle, consonne ou syllabe, *qui varie à chaque nombre selon la personne*, suit le suffixe de conjugaison (baria, nouba) ou le précède (kounama, kanouri, toubou) ou précède le radical verbal (mâba, kodoï). Parfois les aspects négatifs font usage de préfixes (baria), et l'on a dans certaines langues des aspects secondaires à préfixe et d'autres à suffixe. A noter qu'en kanouri et en toubou les aspects secondaires à suffixe placent la syllabe de personne après le suffixe, tandis que les aspects principaux l'intercalent entre le radical verbal et le suffixe. Quoi qu'il en soit, le système de conjugaison présente dans son ensemble une unité réelle et offre des particularités qui constituent la caractéristique principale du groupe.

Ordre des mots : le sujet précède le verbe; le complément d'un verbe précède ce verbe et se place entre le nom sujet

et le verbe ; le complément d'un nom précède ce nom (baria, kounama, nouba) ou bien le suit (kanouri) ou bien le précède ou le suit selon que le nom est indéterminé ou déterminé (toubou, mâba, kodoï, mimi, mara, kondjara, rougna, où l'on dit « maison-maître » pour « un maître de maison » et « maître-maison » pour « le maître de la maison); le qualificatif suit le nom qualifié ; le déterminatif, s'il est adjectif, suit le nom déterminé ou, s'il y a aussi un qualificatif, ce qualificatif; le nom de nombre suit le nom de la chose nombrée, sauf, à ce qu'il semble, en baria. Dans certaines langues (nouba, kanouri, toubou, mâba), on peut suffixer une particule de relation ou d'annexion au complément d'un nom, qu'il précède ou suive ce nom.

L'existence de tons musicaux n'a été signalée dans aucune des langues du groupe.

Il n'y a que des désinences vocaliques en kounama, mais, dans toutes les autres langues, on rencontre des désinences consonantiques.

Le pluriel des noms est indiqué par un suffixe ajouté ou substitué à la désinence du singulier ou par la suppression pure et simple de cette désinence.

Les pronoms personnels se ramènent aux formes essentielles suivantes :

	Kounama	Baria	Nouba	Toubou	Kanouri	Mâba
1re pers. s.	*a, na, ba*	*a, o, i*	*a, ay*	*t, r*	*u, i, sk*	*a*
— pl.	*ma* (excl.), *ka* (incl.)	*be, k, g*	*u*	*t, r*	*d, ye*	*m*
2e pers. s.	*ne, e*	*ç*	*i, ir, n*	*n, m*	*n, m*	*k, m*
— pl.	*me*	*e, k, g*	*u, ur, k*	*n, d*	*n, d, u*	*k*
3e pers. s.	*e, i, u*	*te, u, o*	*ta, tar*	*s, o*	*ši, tši, o*	*t*
— pl.	*o*	*te, k, g*	*te, ter*	*s, d*	*tša, a*	*w*

Le groupe renferme 20 langues aujourd'hui plus ou moins connues, auxquelles il en faut probablement ajouter 10 autres, sur lesquelles nous ne possédons pas de renseignements suffisants pour fixer notre certitude, et peut-être d'autres encore dont nous ignorons même le nom :

1° le **nouba** (*nūba*) ou « nubien » ou *nōb* ou *barabra* ou *berberi*

ou *barabira* ou *boromu* ou *dzoṅ*, parlé le long du Nil depuis Assouan au Nord jusqu'en amont de Khartoum et dans une partie du Sennâr, du Kordofan et du Darfour (beaucoup de « Barbarins » sont domestiques ou petits artisans en Égypte) ; nombreux dialectes : *kenus* ou *kenuz* ou *kenzi* (anciennes provinces des Nobades de Philœ et Talmis, entre Assouan et Ouadi-Ibrim) ; *fadidza* (région de Ouadi-Halfa, entre Ouadi-Ibrim et Dar-Soukkot) ; *mahas* (de Dar-Soukkot à la 3ᵉ cataracte) ; *doṅgola* ou *danagla* (ancien royaume des Nápatai ou Noubádai chrétiens, province de Dongola, de la 3ᵉ à la 5ᵉ cataracte) ; *fundzi* (province de Berber et ancien État de Méroé, Fazogl, Sennâr, Dar-Foundji) ; *n'ima*, *kadero*, *kargo* ou *garko*, *dulmān*, *kolfān* ou *kordofān*, *koldadzi*, *delen* (Nord et Centre du Kordofan) ; *deyer* ou « daïr » ou nouba propre (Dar-Nouba, Sud du Kordofan) ; *midob* (Est du Darfour), etc. (nouba moderne, 1619 Duret et 1650 Carradori) ;

2° le **baria** (*barya*) ou *barea* ou *nere* ou *mogoreb* ou *marda* ou *kolkotta*, parlé à l'Est de l'Atbara, au Nord du Mâreb (1814 Salt) ;

3° le **kounama** (*kunama*) ou *bāza* ou *bāzen* ou *bāden*, parlé dans le Barka, au Nord-Ouest de l'Abyssinie, sur les deux rives du Mâreb et surtout entre le Mâreb et le Takazzé (1814 Salt) ;

4° le **toubou** (*tubu*) ou *tibbu* ou *teda* ou *dāza* ou *gorān*, parlé par les Toubou ou Tibbou ou Téda du Tou ou Tibesti (Bardoa de Léon l'Africain) et du Kaouar, les Boulguéda du Borkou, les Dâza et Kécherda du Bodélé, les Kréda ou Karra du Mourtcha et du Nord-Ouest du Ouadaï, et d'autres tribus englobées avec les précédentes par les Arabes sous le nom de Gorân ou Gourân (1819 Burckhardt) ;

5° à 13° le **takli** ou *takale* ou *tagale* ou *tegele*, le **daghig** (*dagig*) ou *dagig*, le **kadougli** (*kadugli*), le **katla**, le **koalib**, le **krongo**, le **miri**, le **rachad** (*rašad*) et le **tima**, parlés au Kordofan (1916 Meinhof) ;

14° le **kondjara** (*kondzara*) ou *koṅgara* ou *fūr* ou *fōr*, parlé au Darfour dans la région d'El-Facher (1815 Salt) ;

15° le **mara** ou *marra* ou *kanābyaṅ*, parlé par les Mararit ou Abouchârib ou Chouârib et les Ménagon ou Saba de la région

montagneuse située au Sud du Tama, entre le Dar-Massalit à l'Est et le Ouadaï à l'Ouest (1851 Ibn Omar El Tounsy);

16° le **mâba** (*māba*) ou « ouadaïen » ou *kursa* ou *dẓema*, parlé au Ouadaï dans les provinces de Ouara et d'Abécher, ainsi qu'à l'Est et au Sud-Est de ces villes (1817 Vater d'après Seetzen);

17° et 18° le **mimi** (*mīmi*) ou *mīma*, Nord du Ouadaï, et le **kodoï** (*kodoy*), Nord-Est et Sud d'Abécher (1906 Demombynes d'après Decorse);

19° le **rougna** (*ruña*) ou *ruṅga* ou *ruń'a*, parlé par les païens dits Fertit ou Djénakhéra du Dar-Rougna ou Dar-Rounga et du Dar-Kouti, au Sud du Darfour et du Ouadaï (1849 Clarke);

20° le **kanouri** (*kanuri*) ou *bọrnu* ou *barnu* ou *baribari* ou *balibali* ou *aẓa* ou *kaga* ou *kagatsan* ou *ẓanẓanti* ou *bino* ou *mafak* ou *kaniki*, etc., parlé dans le Kanem, le Bornou, le Manga, le Mounio, le Damergou, etc., par plus d'un million d'individus (1819 Burckhardt et Bowdich);

21° à 30° le **zaghaoua** (*ẓagāwa*) ou *ẓọgāwa* ou *berri* ou *bela* ou *bele* ou *bede* ou *bedeyāt* ou *wań'a*, parlé au Sud du désert de Libye et au Nord du Darfour et du Ouadaï, notamment dans la région du Kabga ou Gaoga de Léon l'Africain (1912 MacMichael); — l'**anna**, parlé dans l'Ennedi (néant); — le **siga** ou *sigato* ou *sulgu* ou *kurmu* ou *bayti* ou *berti*, dans l'Est du Darfour, au Sud des Nouba du Midob (1922 MacMichael); — le **tama**, dans le Dar-Tama, au Nord-Est du Ouadaï, parlé aussi, concurremment avec l'arabe, par les Kimr, qui habitent au Nord des Massalit et à l'Est des Tama (néant); — le **massalit**, entre le Darfour et le Ouadaï, dialectes *ereṅga* et *mūn* (1922 MacMichael); — le **soungor** (*suṅgor*) ou *murro*, au Sud du précédent (néant); — le **kachméré** (*kašmere*), Sud du Ouadaï (1906 Demombynes); — le **bigna** (*biña*) ou *miṅga*, à l'Est du Dar-Rougna (néant); — le **dadio** (*dad'o*) ou *dadẓo* ou *dadẓo* ou *dāgo* ou *tāgu* ou *finiṅga*, parlé au Sud du Ouadaï dans les régions d'Amdam, de Goz-Beïda et du Dar-Sila, dans la haute vallée de la Batha et celle du Bahr-Salamat, et aussi dans le Sud du Darfour et au Kordofan (dialecte *bego* ou *beo* ou *bayko*) (1922 MacMichael); — le **kadiaksé** (*kad'akse*), ou *kadẓarge* ou *kadẓara* ou *kagaru* ou *murgi* ou *birked*, province de Goz-Beïda et Darfour (1922 MacMichael).

BIBLIOGRAPHIE. — Aucun ouvrage ne traite de l'ensemble du groupe. Parmi les publications concernant les langues étudiées spécialement, on peut citer : pour le nouba ancien, GRIFFITH, *The Nubian texts of the Christian period*, Berlin, 1914, in-8 ; pour le nouba moderne, Leo REINISCH, *Die Nuba-Sprache*, Wien, 1879, 2 vol. in-8; et surtout R. LEPSIUS, *Nubische Grammatik*, Berlin, 1880, in-8; pour le baria et pour le kounama, Leo REINISCH, *Die Barea-Sprache*, Wien, 1874, in-8, et *Die Kunama-Sprache*, Wien, 1888-1891, in-8; pour le toubou, CARBOU, *La région du Tchad et du Ouadaï*, Paris, 1912, 2 vol. in-8 (1er vol.); pour le mâba, BARTH, *Sammlung und Bearbeitung Central-Afrikanischer Vokabularien*, Gotha, 1862-1866, 2 vol. in-8 (chap. XII); pour le kanouri, A. VON DUISBURG, *Grundriss der Kanuri-Sprache*, Berlin, 1913, in-8, et surtout Dr. NOEL, *Petit manuel français-kanouri*, Paris, 1923, in-12.

II. — *Groupe nilo-abyssinien* (15 langues).

Ce groupe occupe la vallée du moyen Nil Bleu et celle du Nil Blanc (Bahr-el-Abiod et Bahr-el-Djebel) depuis la hauteur de Sennâr en aval jusque près de Lado en amont, s'étendant à l'Est le long du Sobat et jusque sur le bas Omo et à l'Ouest sur le bas cours du Bahr-el-Arab et du Soueh, avec deux enclaves au Sud, l'une sur le haut Bahr-el-Djebel à hauteur de Wadelaï et l'autre à l'Est du lac Victoria. Les langues qui le composent sont parlées par un nombre assez considérable de nègres purs dont la civilisation semble en général très arriérée.

Les vestiges de classes nominales sont bien apparents : des préfixes de classe ont subsisté fréquemment (*o*, *a* et *u* en chilouk, *a* en aniouak, *ẹ* et *kẹ* en nouër, *a*, *e*, *o*, *u* et *ke* en dinka, etc.); le qualificatif est souvent précédé d'un ancien pronom de classe (*ma* ou *mẹ* en chilouk, *ma* ou *mu* en aniouak, *me* en nouër); d'anciens pronoms de classe se suffixent au nom pour en marquer le nombre ou la détermination ou s'intercalent entre le nom et son complément (dinka : *ran* « homme », *ran e* « l'homme »; *tyam* « manger », *tyam de* « le manger »; *rōr* « des hommes », *rōr ke* « les hommes »).

On a cherché à distinguer le nom du verbe en maintenant certains préfixes nominaux et en usant de désinences spéciale-

ment nominales, mais, dans beaucoup de cas, la distinction repose uniquement sur la place donnée respectivement au nom et au verbe.

Les aspects principaux du verbe sont indiqués tantôt par des suffixes (chilouk), tantôt par des suffixes pour une première forme et des préfixes pour une deuxième (aniouak), tantôt par des suffixes à certains aspects et des préfixes à d'autres (nouër), tantôt par des préfixes (dinka). La forme du suffixe ou du préfixe servant à marquer un aspect donné est analogue dans la plupart des langues du groupe (aoriste : suffixe *ϱ* en chilouk et en aniouak ; parfait : suffixe *i* en chilouk, suffixe *i* ou préfixe *tše* en aniouak, préfixe *tši* en nouër, *k'i* en dinka, *k* en mékan). Lés aspects négatifs ou secondaires sont indiqués généralement par des préfixes. Lorsqu'il est fait usage de préfixes, ceux-ci se préfixent souvent, non pas au radical verbal, mais au pronom sujet, s'il existe.

Ordre des mots : le sujet précède le verbe, sauf parfois en nouër et, pour ce qui est du pronom sujet, à l'injonctif ; le complément d'un verbe peut suivre ce verbe (dinka, parfois chilouk), mais le précède en général (mékan, aniouak, nouër, gamila, berta) ; le complément d'un nom suit ce nom ; le qualificatif suit le nom qualifié ; le déterminatif suit le nom déterminé ou, s'il y a aussi qualificatif, ce qualificatif, sauf en ce qui concerne le déterminatif jouant le rôle de notre article défini qui, lui, se suffixe toujours directement au nom ; le nom de nombre suit le nom de la chose nombrée.

Il existe des tons musicaux à valeur étymologique et à valeur grammaticale, souvent employés pour distinguer le pluriel du singulier dans les noms et le passif de l'actif dans les verbes. Les désinences consonantiques sont très fréquentes.

Le pluriel des noms est indiqué, dans chaque langue du groupe, au moyen de divers procédés employés soit isolément, soit simultanément : chute ou modification de la désinence du singulier ou addition au radical d'une désinence spéciale ; changement apporté à la quantité, à la nature ou à la tonalité de la voyelle radicale ; modification de la dernière consonne radicale,

la sonore devenant sourde ; changement ou addition de préfixe de classe.

Les pronoms personnels sont les suivants dans les langues les mieux connues :

	Chilouk	Aniouak	Nouër	Dinka
1re pers. sing.	*ya, a*	*a*	*a*	*ya, a*
— plur.	*wa, wọ*	*wa*	*kọ*	*o, wa*
2e pers. sing.	*yi, i*	*i*	*i, d'i*	*yi, i, u*
— plur.	*wu*	*wu*	*yę, u*	*we, un*
3e pers. sing.	*yę, ę, gọ*	*ę*	*ę, d'ę*	*ye, e*
— plur.	*gę*	*gi*	*kę, k'ę*	*ke, en*

Le groupe renferme 8 langues aujourd'hui à peu près bien connues, auxquelles il faut probablement en ajouter 7 autres sur lesquelles nous ne possédons pas de renseignements suffisants et peut-être d'autres encore que nous ignorons totalement aujourd'hui :

1° le **chilouk** (*šiluk*) ou *šilluk* ou *tsolo* ou *dembo*, parlé sur la rive gauche du Nil Blanc au Nord de Fachoda et sur les deux rives entre Fachoda et le confluent du Sobat, ainsi que, plus au Sud, entre le Bahr-el-Djebel et le bas Omo (1829 Rüppell) ;

2° le **nouër** (*nuęr*) ou *nāþ* ou *abigar*, parlé dans la région marécageuse qui borde les deux rives du bas Bahr-el-Djebel et du bas Sobat, au Sud du précédent, ainsi que le long du moyen Sobat (1861 Brun-Rollet) ;

3° l'**aniouak** (*an'wak*) ou *yambo*, parlé à l'Est du nouër et du dinka dans la vallée du Sobat, ainsi que dans celle du Baro et entre les rivières Gélo et Akobo (1872 d'Abbadie) ; peut-être faut-il lui rattacher le *kogo*, parlé sur la rive droite du Baro ;

4° le **dinka** (*diṅka*) ou *deṅka* ou *dźeṅg* ou *d'aṅg* ou *d'eṅ* (plur. *diṅk*), parlé entre le Nil Blanc (Bahr-el-Abiod) et le Nil Bleu, que les Dinka déborderaient vers l'Est dans la direction du lac Tsana et, en amont des Nouër, sur les deux rives du Sobat, sur la rive droite du Bahr-el-Djebel (tribu des Bor) et sur sa rive gauche (tribus des Gog, des Agar, des Lessi, des Atot, des Eliab, etc.) (1829 Rüppell) ; peut-être faut-il y rattacher le *gindilia*

(entre le bas Sobat et le coude du Nil Bleu), ainsi que le *barun*, le *komo* et le *mao*, parlés entre le Baro et le même coude;

5° le **diour** (*d'ūr*) ou *dẓūr* ou *luō* ou *ber*, parlé sur le bas Soueh (1872 Schweinfurth);

6° le **gang** (*gañ*) ou *gani* ou *atšoli* ou *šuli* ou *šefalu* ou *obbo* ou *faradẓoke*, sur la rive droite du haut Bahr-el-Djebel, entre Wadelaï et le lac Albert (1882 Emin Bey);

7° le **lour** (*lūr*) ou *lūri* ou *alūru* ou *dūdu* ou *alūlu*, sur la rive gauche, en face du précédent (1882 Emin Bey);

8° le **djalouo** (*dẓaluo*), à l'Est du lac Victoria, au Sud des Kavirondo (1904 Johnston);

9° à 15° sous réserves : le **tabi**, sur le Nil Bleu, au Sud-Est de Khartoum, en voie de disparition (1864 Marno et 1912 Seligmann); — le **goulé** (*gule*), dans le Dar-Foundji et le Fazogl, au Sud du Sennâr (1864 Marno et 1912 Seligmann); — le **hamedj** (*hamedẓ*) ou *naga* (ce dernier nom, qui semble d'ordre géographique, est donné aussi au « gounza », langue probablement chamitique parlée dans le voisinage des Hamedj), parlé en aval du coude du Nil Bleu (1864 Marno); — le **berta** ou *mugo* ou *dalla* ou *doba*, parlé par les Beni Changoul, entre le Nil Bleu et le Nil Blanc, au Sud de Sennâr (1864 Marno et 1874 Halévy); — le **gamila** ou *gamolla*, parlé par une fraction des Berta, dite « Qamâmil » par les Arabes, sur le Dabous (1826 Cailliaud); — le **mékan** (*mekan*) ou *sūro* (*sūro* en kaffa signifie « nègre »), parlé au Sud-Ouest du Kaffa, entre le haut Akobo et l'Omo, et qui semble tenir à la fois du groupe nilo-abyssinien et, à un degré moindre, du groupe nilo-équatorien (1913 Conti-Rossini); — le **doko** ou *golda* ou *madẓi* ou *maše* ou *mase*, sur les deux rives du bas Omo (1913 Conti-Rossini).

BIBLIOGRAPHIE. — Aucun ouvrage ne traite de l'ensemble du groupe. Parmi les publications concernant spécialement les langues les mieux étudiées, on peut citer : pour le chilouk, D. WESTERMANN, *A short grammar of the Shilluk language*, Philadelphia, 1911, in-16 ; pour le nouër, même auteur, *The Nuer language*, Berlin, 1912, in-8 ; pour l'aniouak, même auteur, *Some notes and a short vocabulary of the Anywak language*, Berlin, 1912, in-8 ; pour le dinka, G. BELTRAME, *Grammatica e vocabolario della lingua Denka*, Firenze, 1881.

(2e édition), in-8; pour le gang, A. L. KITCHING, *An outline of the Gang language*, London, 1907, pet. in-8.

III. — *Groupe nilo-équatorien* (26 langues).

Ce groupe fait suite, dans la direction du Sud, au groupe nilo-abyssinien, dont il entoure les enclaves méridionales, son domaine s'entremêlant à l'Ouest avec celui du groupe nilo-congolais. Il commence, au Nord, à peu près à hauteur de Lado, pour s'étendre, au Sud, jusqu'au 6° de latitude sud ; à l'Ouest, il dépasse le Bahr-el-Djebel et, à l'Est, il atteint le lac Stéphanie et le mont Kilimandjaro, poussant même une antenne jusqu'au rivage de l'Océan Indien près de Mombassa.

Les langues de ce groupe sont parlées à la fois par des populations agricoles de race nègre (Bari, Kouafi, etc.), par des populations pastorales, beaucoup moins nombreuses, dont certaines fractions semblent se rattacher en partie à un rameau de la race blanche (Massaï propres, Houmba, Toussi, Nkolé) et enfin par quelques castes de chasseurs et d'artisans dans lesquelles l'élément nègre coudoie l'élément négrille (Dorobo, Koünono, etc.). L'ensemble constitue un groupe ethnique très hétérogène, de civilisations plutôt arriérées. Il semble que les éléments de race blanche qui en font partie aient, une fois installés dans le pays, emprunté le parler des nègres qui s'y trouvaient avant eux.

Les classes nominales sont très nettes. Dans chaque langue, il en existe une comprenant les noms des êtres mâles ou forts, les augmentatifs, le nom de la terre et celui de la lune, les noms de collectivités humaines (métiers, tribus, etc.) et des noms divers, puis une autre comprenant les noms des êtres femelles ou faibles, les diminutifs, le nom du ciel et celui du soleil et des noms divers ; en plus de ces deux classes principales, il s'en rencontre souvent d'autres (par exemple, en massaï, une classe des noms de lieux et une classe d'abstractions). Chaque classe est caractérisée grammaticalement : par un préfixe de classe qu'on a improprement qualifié d'article (en bari et en massaï, l'élément

essentiel est *l* pour la 1[re] classe, *n* pour la 2[e]); par un pronom de classe analogue au préfixe et employé devant le complément d'un nom pour remplacer ou rappeler celui-ci; par un relatif de classe (1[re] classe, bari *lo*, massaï *o*; 2[e] classe, bari et massaï *na*) qui se préfixe à l'adjectif pour marquer son accord en classe avec le nom qualifié; enfin par l'emploi, pour chaque classe, d'une série spéciale de déterminatifs ayant pour base le préfixe ou le pronom de classe. Parfois cependant, il y a disparition fréquente du préfixe nominal de classe, comme en bari, ou bien tendance à faire usage d'un pronom personnel commun à toutes les classes, comme en massaï; c'est surtout au pluriel que cette tendance à l'unification des pronoms de classe est manifeste.

Dans les langues, comme le bari, où le préfixe nominal de classe a souvent disparu, on distingue le nom du verbe à l'aide de désinences spécialement nominales, variant avec le nombre; ce procédé se rencontre même dans des langues qui, comme le massaï, ont conservé les préfixes de classe.

Les aspects du verbe sont marqués en général par des suffixes à la voix affirmative et par des préfixes à la voix négative. De plus, on fait usage, à certaines personnes de certains temps, du redoublement parfait ou imparfait du radical. Il existe un passif, indiqué par des suffixes spéciaux.

Ordre des mots : le sujet précède le verbe dans les propositions principales, mais le suit en général dans les propositions relatives lorsque ce n'est pas le relatif qui est sujet (toutefois, en massaï, le nom sujet suit toujours le verbe,. mais il est représenté devant celui-ci par un pronom à forme abrégée); le complément d'un verbe suit ce verbe et se place généralement après le sujet quand celui-ci est suffixé au verbe; le complément d'un nom suit ce nom et est précédé le plus souvent du pronom de classe de celui-ci; le qualificatif suit le nom qualifié et est précédé en général du relatif de classe de ce nom (quelquefois il précède le nom, qui perd alors son préfixe de classe pour le passer au qualificatif); le déterminatif *précède* le nom déterminé, qui perd son préfixe de classe; le nom de nombre est un adjectif et se comporte comme le qualificatif.

Il existe des tons musicaux, dont il est fait usage pour préciser le rôle d'un mot lorsque sa place dans la phrase ne suffit pas à l'indiquer nettement : ainsi, en massaï, le ton haut marque qu'un nom est sujet, le ton bas qu'il est complément d'un verbe.

Le passage d'un radical d'une classe nominale à une autre, ou le changement d'acception dans une même classe, peut amener la modification d'une consonne radicale (bari : *lu-ñatšer* « frère », *ki-yatšer* « sœur » ; massaï *ol-alatše* « frère », *eñ-anatše* « sœur » ; *eñ-geray* « enfant », *eñ-gelay* « petit enfant »). Les désinences consonantiques sont fréquentes.

Le pluriel des noms est indiqué à la fois par une modification du préfixe de classe et par une modification de la désinence, ce dernier procédé étant naturellement le seul employé là où le préfixe de classe a disparu.

Voici les pronoms personnels en bari et en massaï :

	Bari	Massaï
1re pers. sing.	*nan*	*na, a*
— plur.	*i*	*yok, ki*
2e pers. sing.	*do, da*	*ye, i*
— plur.	*ta*	*nday, i*
3e pers. sing. 1re classe	*lo, lu, la*	*ol, o*
— — 2e classe	*na, nu, no*	*en, na*
— — 3e classe	......	*ne*
— — 4e classe	......	*ki*
— — commun	......	*ni, e*
— plur. commun	*tše, k'e*	*nd'e, e*

Le groupe renferme 15 langues aujourd'hui plus ou moins bien connues, auxquelles il en faut probablement ajouter 11 autres sur lesquelles nous sommes insuffisamment renseignés pour nous prononcer à coup sûr et peut-être d'autres encore dont nous ignorons le nom :

1° le **bari** ou *džilio*, parlé par 300.000 individus environ des deux côtés du Bahr-el-Djebel entre Lado et Doufilé et, vers l'Est, jusque près du lac Rodolphe : dialectes *bari* propre, *fad'ellu*, *liggi* ou *yeyi* ou *yeh*, *mandari*, *šir* ou *k'ir*, *kuko*, etc. (1864 F. Müller);

2° le **latouka** (*latuka*), à l'Est du Bahr-el-Djebel, entre les 5° et 4° de latitude nord (1867 Baker);

3° et 4° le **tourkana** (*turkana*) ou *kume*, sur la rive occidentale du lac Rodolphe, et le **souk** (*suk*), au Sud-Est et à l'Est du même lac, parlé peut-être même jusqu'au Nord-Est du lac Stéphanie sous le nom de *konso* (1902 Hobley);

5° le **liri** ou *riri* ou *lira* ou *lir* ou *sugu*, à l'Est de la pointe septentrionale du lac Albert (1852 H. C. von der Gabelentz);

6° le **nandi** ou *lumbwa* ou *teñwal*, au Nord-Est du lac Victoria (1902 Hobley);

7° le **kavirondo** ou *kapirondo*, à l'Est et au Sud-Est du lac Victoria, le long des rives du lac et du bas Nyando (1873 New);

8° le **kipsikissi** (*kipsikisi*) ou *kakesan*, appelé aussi *lumbwa* comme le nandi, au Sud-Est du lac Victoria, en arrière du Kavirondo (1902 Hobley);

9° le **houmba** (*humba*) ou *wamba* ou *hima*, parlé par des pasteurs négroïdes vraisemblablement apparentés à la race blanche, dits Vahoumba, Vahimba, Vahouma, Vahima [1], etc., éparpillés entre le lac Albert et le lac Édouard dans l'Ounioro et entre les lacs Victoria et Tanganika, ainsi qu'au Sud-Est du lac Victoria (1885 Last);

10° le **toussi** (*tusi*), parlé par des pasteurs de même origine que les précédents, éparpillés au Sud du lac Victoria dans l'Ouniamouézi et entre les lacs Victoria et Tanganika dans l'Ouha (1907 Dahl);

11° le **tatourou** (*taturu*), au Sud-Est du lac Victoria, entre le pays des Kipsikissi et le lac Niarassa, à l'Est de l'Ouniamouézi (1885 Last);

12° le **massaï** propre (*mâsai*), parlé seulement par 30.000 à 40.000 pasteurs négroïdes se donnant eux-mêmes le nom de *Il-*

1. D'autres populations négroïdes pastorales, désignées par les Bantou sous le même nom de *Vahima*, *Bahima*, *Wašimba*, etc., se rencontrent plus à l'Ouest jusque près de l'Océan Atlantique. Elles passent pour parler les diverses langues bantou sur les territoires desquelles elles nomadisent. Il est possible que les Hottentots soient issus d'un mélange de ces mêmes « Vahima » avec les négrilles dits « Bushmen » qui les avoisinent et dont ils auraient adopté la langue.

māsa-e (sing. *Ol-māsa-ni*), répandus surtout entre le lac Victoria et le mont Kénia (1857 Erhardt);

14° le **dorobo** ou *ndorobbo* ou *torōbo*, parlé par des castes de chasseurs et d'artisans disséminés parmi les Massaï et les Kouafi et ne se distinguant que très peu du massaï propre et du kouafi (1904 Johnston);

14° le **kouafi** (*kwafi*) ou *kwavi* ou *kwapi* ou *loykop* ou *loygob* ou *kisōgo*, appelé aussi *lumbwa* comme le nandi et le kipsikissi par les Massaï, parlé par un nombre considérable de noirs sédentaires et agriculteurs habitant au Nord, à l'Ouest et au Sud du mont Kilimandjaro, du 1° au 6° de latitude sud, et confondus souvent avec leurs voisins nomades et pasteurs du Nord sous le nom de Massaï (1854 Krapf);

15° le **mbouloungué** (*mbulunge*) ou *buruñgi*, au Sud-Ouest du précédent, entre les 5° et 6° de latitude sud (1906 Meinhof);

16° à 26° sous réserves : le **karamodjo** (*karamodžo*) ou *kakisera*, à l'Ouest du tourkana et du souk; le **kamassia** (*kamas'a*), au Sud du précédent ; le **lango** ou *kedi* ou *kidi* ou *lukedi*, au Nord du lac Victoria, à l'Est de l'Ounioro ; le **nguichou** (*ñgišu*), au Sud-Est du mont Elgon, au Nord du nandi ; le **goumi** (*gumi*) ou *elgumi*, au voisinage des Massaï propres (pour ces cinq langues, 1904 Johnston); — le **nkolé** (*ñkole*), langue de pasteurs répandus entre le lac Édouard et le lac Victoria; le **nifoua** (*nifwa*) et le **gaya**, sur la côte orientale du lac Victoria (pour ces trois langues, néant) ; — le **tatoga**, région montagneuse du Nord de l'ancien Est-Africain allemand (1898 Werther et 1911 Struck) ; — l'**irakou** (*iraku*), même région (1911 Struck) ; — le **mbougou** (*mbugu*), région de Mombassa (1906 Meinhof).

BIBLIOGRAPHIE. — Aucun ouvrage ne traite de l'ensemble du groupe. Parmi les publications concernant spécialement les langues les mieux étudiées, on peut citer : pour le bari, J. C. MITTERUTZNER, *Die Sprache der Bari*, Brixen, 1867, in-8, et traduction anglaise par C. R. OWEN (*Bari grammar and vocabulary*, Oxford, 1908, in-8) ; pour le liri, L. WOLF, *Beitrag zur Ki-Lir-Sprache (Sugu)*, dans *Zeitschrift für Afrikanische Sprachen*, III; pour le souk, W. H. MERVYN BEECH, *The Suk, their language and folklore*, Oxford, 1911, in-8 ; pour le nandi, A. C. HOLLIS, *The Nandi, their language and folklore*, Oxford, 1909, in-8 ;

pour le kavirondo, M. Wakefield, *Vocabulary of the Kavirondo language*, London, 1887, in-8; pour le massaï propre, A. C. Hollis, *The Masai, their language and folklore*, Oxford, 1905, in-8; pour le kouafi, F. Müller, *Die Sprache der Il-Oigob*, Wien, 1884, in-8.

IV. — *Groupe kordofanien* (10 langues).

Ce groupe constitue géographiquement une sorte de tampon isolateur entre le groupe nilo-tchadien et le groupe nilo-abyssinien. Il est limité à la partie sud-occidentale du Kordofan et comprend, d'après le peu que nous en savons, dix langues qui sont parlées chacune, sur une aire très restreinte, par des nègres qui diffèrent à la fois des Nouba et des Chilouk, leurs voisins les plus immédiats.

Le système linguistique des parlers de ce groupe correspond très étroitement à celui des langues bantou et les classes nominales s'y présentent sous le même aspect que dans ces dernières. Chaque classe (en talodi, il en existe dix-sept, dont dix pour le singulier et sept pour le pluriel) est caractérisée par un pronom de classe spécial qui se préfixe à tout nom de sa classe et à l'adjectif qualifiant celui-ci et qui sert en outre de pronom personnel, de relatif, de démonstratif et de déterminatif (en talodi avec la voyelle *a* pour marquer la proximité, la voyelle *i* ou *u* pour marquer l'éloignement).

Le préfixe de classe suffit à distinguer le nom du verbe; la place respective de l'un et de l'autre complète la distinction.

Les documents existants ne permettent pas de préciser quelle est exactement la conjugaison du verbe.

Ordre des mots : le nom sujet précède le verbe; le pronom sujet le précède également dans les propositions principales, mais le suit en général dans les propositions relatives, et alors la consonne initiale de la racine verbale passe de l'occlusive à la constrictive correspondante (comparez la modification inverse en peul); le complément d'un verbe suit ce verbe; le complément d'un nom suit ce nom, avec ou sans particule de relation interca-

laire (*n'* en talodi) ; le qualificatif, le nom de nombre et le déterminatif suivent le nom qualifié, nombré ou déterminé.

L'existence de tons musicaux n'a pas été signalée.

Les désinences consonantiques sont fort rares et peut-être ne se rencontrent-elles qu'autant que la consonne terminale est une sonante.

Le pluriel des noms est indiqué par le changement du préfixe de classe. Quelquefois on ajoute au radical un suffixe de pluralité.

Enfin il semble que la consonne initiale de la racine peut subir des modifications selon le nombre ou la classe.

En talodi, les pronoms personnels apparaissent comme suit :

1re pers. sing. *i*, *ya*, *g'a*, plur. *ri*, *r'a*

2e pers. sing. *u*, *o*, *ba*, *wa*, plur. *ta*, *da*

3e pers. (pronom de classe).

Le groupe comprend, d'après Meinhof, les 10 langues suivantes : **talodi**, **eliri**, **lafofa** ou *lifofa*, **toumtoum** (*tumtum*), **kanderma** ou *kinderma*, **kaouama** (*kawama*) ou *kawarma*, **loumoun** (*lumun*), **cheïboun** (*šeybun*), **tagoï** (*tagoy*) et **toumèli** (*tumẹli*) ou *tumale*.

BIBLIOGRAPHIE. — Le toumèli a été étudié dès 1848 par L. TUTSCHEK (*Ueber die Tumale-Sprache*, München, in-4). Mais c'est seulement en 1916 que Carl MEINHOF (*Eine Studienfahrt nach Kordofan*, dans *Abhandlungen des Hamburgischen Kolonialinstituts*, Band XXXV) l'a séparé du takli et autres langues nilo-tchadiennes avec lesquelles on l'avait d'abord rangé, pour le classer avec le talodi et les autres langues kordofaniennes. Tout ce que nous savons de celles-ci est consigné dans deux communications de la *Zeitschrift für Kolonial-Sprachen*, l'une de Mme Brenda Z. SELIGMANN (*Notes on the languages of the tribes of Southern Kordofan*, 1910-1911), l'autre de Ferdinand BORK (*Zu den neuen Sprachen von Süd-Kordofan*, 1912-1913) et dans le travail susmentionné de Meinhof ; ce dernier est le seul qui ait parlé du cheïboun et du tagoï.

V. — *Groupe nilo-congolais* (19 langues).

L'aire territoriale du groupe nilo-congolais, assez étroite, s'insinue entre celles du groupe nilo-abyssinien au Nord-Est, du groupe nilo-équatorien à l'Est, du groupe bantou et d'une antenne

du groupe oubanguien au Sud et de ce dernier groupe à l'Ouest et au Nord-Ouest. Dans l'ordre linguistique, il offre de nombreux points de contact avec les divers groupes dont il est entouré. Les populations parlant les langues dont il se compose appartiennent toutes à la race nègre; elles ne paraissent pas numériquement considérables et ont une civilisation en général arriérée.

Le système des classes nominales se révèle par des préfixes de classe accolés à la plupart des noms et peut-être, quoique ce ne soit pas sûr, par l'emploi de certains de ces préfixes devant l'adjectif. Il ne semble pas être fait usage actuellement de pronoms de classe.

L'emploi, très net encore, de préfixes nominaux concourt à faire distinguer le nom du verbe; la distinction est assurée en outre par la place respective de l'un et de l'autre.

Les documents que nous possédons ne permettent pas de déterminer exactement les procédés de conjugaison.

Quant à l'ordre des mots, il se différencie nettement de ceux qui sont observés dans tous les groupes géographiquement voisins, en ce sens que le complément d'un nom précède ce nom, au lieu de le suivre comme dans les langues nilo-abyssiniennes, nilo-équatoriennes, bantou, oubanguiennes et kordofaniennes (mangbètou : *nẹ-ñgo* « œil », *ke-pi* « peau », *nẹ-ñgo ke-pi* « peau d'œil, paupière »; mbouba : *o-yo* « œil », *ku-pu* « peau », *o-yo pu* « paupière » ; léga : *e-tsa* « maison », *lẹ-lio* « bouche », *e-tsa lio* « bouche de maison, porte »; lendou : *tsa* « maison », *la-tso* « bouche », *tsa tso* « porte »). Le nom complété garde quelquefois, mais perd, généralement, son préfixe de classe (voir les exemples cités plus haut). Comme dans les langues oubanguiennes, le sujet précède le verbe, le complément d'un verbe le suit. Le déterminatif ou le nom de nombre se place après le nom déterminé ou nombré. Le qualificatif paraît également suivre le nom qualifié.

L'existence de tons musicaux semble douteuse.

On ne rencontre pas de désinences consonantiques; lorsqu'une racine se termine par une consonne, on lui suffixe, pour en faire

un mot, une voyelle qui est généralement identique ou analogue à la voyelle radicale (mangbètou : $\sqrt{wur}$ donne *no-wuru* « montagne », $\sqrt{pañg}$ donne *mu-pañga* « lance » ; aouidi : $\sqrt{lol}$ donne *lolo* « mâle » ; mbouba : $\sqrt{tseb}$ donne *ge-tsebe* « terre » ; léga : $\sqrt{sogb}$ donne *i-sogbo* « terre »). Il semble d'ailleurs que le plus grand nombre des racines affecte la forme « consonne + voyelle ».

Le pluriel des noms est marqué par le changement du préfixe de classe (mangbètou : *na-bẹ* « homme », plur. *a-bẹ*).

Le pronom de la 1re pers. du sing. est *ma* en mangbètou, en aouidi, en madi, en bangba, etc., et celui de la 1re pers. du plur. *a-ma* ; ceux de la 2e et de la 3e pers. sont respectivement en mangbètou *mi* et *nẹ* au sing., *a-mi* et *a-ẹ* ou *ẹ* au plur. A noter la correspondance du préfixe de classe mangbètou *nẹ* ou *na* (plur. *ẹ* ou *a*), qui semble le plus fréquent, avec les pronoms *nẹ* et *a-ẹ* ou *ẹ* (comparez les préfixes bangba *na*, léga *lẹ*, lendou *la*, correspondant à mangbètou *nẹ* ou *na*).

L'état actuel de notre documentation permet d'attribuer au groupe nilo-congolais 19 langues, mais il est possible qu'il en contienne d'autres, encore inconnues :

1° l'**aouidi** (*awidi*), parlé par une petite tribu habitant sur la rive orientale du Bahr-el-Djebel, à hauteur de Doufilé, à laquelle on donne parfois le même nom d'Aboukaya qu'à une fraction des Mittou (1863 et 1865 Miani) ;

2° le **morou** (*moru*) ou *amadi* (ne pas confondre avec le mittou, dit aussi « amadi », du groupe oubanguien), parlé dans le village et la région d'Amadi-sur-Yeï, au Nord-Ouest de Lado (1892 Casati) ;

3° et 4° sous réserves : le **niangbara** (*n'ãgbara*) ou *n'ambara*, parlé au Sud du précédent sur la rivière Yeï ou Djeï (1867 Mitterutzner d'après Morlang) ; — le **kédérou** (*kederu*), à l'Ouest du précédent (1880 Wilson et Felkin) ;

5° le **madi** ou *tšopi* (ne pas confondre avec le morou, dit « amadi », qui est du même groupe, ni surtout avec le mittou, dit aussi « amadi », du groupe oubanguien), sur la rive orientale

du Bahr-el-Djebel, au Sud des Aouidi, entre Doufilé et Wadelaï (1880 Wilson et Felkin);

6° à 8° le **logbouari** (*logbwari*) ou *lugware* ou *mogwari*, dans le district montagneux d'où sortent, vers le Nord, la Yeï, vers l'Est des affluents du haut Bahr-el-Djebel et vers l'Ouest des affluents du haut Ouellé; le **lendou** (*lendu*) ou *alendu*, au Sud du précédent, à l'Ouest et à hauteur de Wadelaï, près des sources du Kibali ou haut Ouellé, avec une enclave dans l'Ounioro; le **léga** (*lega*) ou *legga* ou *drugu*, au Sud du précédent, dans la direction du lac Albert (pour ces trois langues, 1904 Johnston);

9° le **logo** ou *loggo*, à l'Ouest du logbouari, entre le Doungou et le Kibali, à l'Est de leur confluent (1921 de Calonne-Beaufaict);

10° le **mangbètou** (*mãgbẹtu*) ou *mãgbatu* ou *mambẹktu* ou *mambettu* ou *mombuttu*, parlé à cheval sur le Kibali et au Sud jusqu'au haut Itouri ou Arouwimi (mangbètou de l'Est), puis, plus à l'Ouest, chez les pêcheurs dits Bakango, entre l'Ouellé et son affluent le Bomokandi (mangbètou du centre), et, plus à l'Ouest encore, entre l'Ouellé et l'Arouwimi (mangbètou de l'Ouest ou *makẹrẹ*), langue d'une certaine extension (1876 Long);

11° à 19° sous réserves : le **madjé** (*madže*) ou *medže*, dans le voisinage et au Sud des Mangbètou de l'Est (néant); — le **momvou** (*momvu*) ou *momwu* ou *momu* ou *mõfu*, au Sud du Kibali, entre les Mangbètou de l'Est et ceux du centre (1888 Junker); — le **lessé** (*lese*), parlé par les *Balese* ou *Walese* à l'Ouest des Momvou et au Sud-Est des Mangbètou du centre (1910 Struck); — le **karè** (*karẹ*) ou *akarẹ* ou *akalẹ* ou *ap'a*, parlé dans quelques colonies, que d'aucuns disent d'origine bantou, éparses parmi les Zandé, entre le Mbomou et l'Ouellé (1888 Junker); — le **bangba** (*bãgba*) ou *bãmba* ou *abãgba*, sur la rive gauche de l'Ouellé, au Nord des Mangbètou du centre (1888 Junker); — le **birri** ou *abiri* ou *ambili*, au voisinage du précédent (1888 Junker); — le **baboua** (*babwa*) ou *ababwa*, sur la même rive, au Nord des Mangbètou de l'Ouest (néant) [1]; — le **bouté** (*bute*) ou

1. D'après le Père Van den Plas, le birri et le baboua seraient des langues

buti ou *mbute* ou *bambute*, sur le haut Itouri, au Sud des Mangbètou de l'Est (1904 Johnston); — le **mbouba** (*mbuba*) ou *boko* ou *buku* ou *koko* ou *kiboko* ou *bukoko*, à l'Ouest du Semliki et à l'Est entre le lac Albert et le lac Édouard, parmi des Bantou du Rouwenzori (1904 Johnston).

BIBLIOGRAPHIE. — Il n'existe aucun ouvrage d'ensemble sur le groupe nilo-congolais ni aucune étude sérieuse de l'une quelconque de ses langues et aucun texte n'a été recueilli. Nous ne possédons que quelques petits vocabulaires de valeur inégale, dont les principaux sont ceux qui ont été donnés par MIANI (aouidi), *Lingua degli Auidi*, dans *Il commercio d'Egitto*, Cairo, settembre 1863; par CASATI (morou, mangbètou, bangba), *Dix années en Equatoria*, Paris, 1862, in-4; par JUNKER (mangbètou, momvou, karè, bangba, birri), *Verzeichnis von Wörtern Zentralafrikanischer Sprachen*, dans *Zeitschrift für Afrikanische Sprachen*, 1888, et d'après lui par F. MÜLLER, *Die æquatoriale Sprachfamilie*, Wien, 1889, in-8; par WILSON et FELKIN (kédérou, madi) *U-Ganda*, London, 1880, in-8; par Sir H. JOHNSTON (madi, logbouari, lendou, léga, mbouba, bouté), *The Uganda Protectorate*, London, 1902-1904, 2 vol. gr. in-8, et enfin par B. STRUCK, d'après Emin Pacha (lendou, léga, mbouba), dans *Mitteilungen des Seminars für Orientalische Sprachen*, XIII, 3, 1910.

VI. — *Groupe oubanguien* (25 langues).

Ce groupe se soude au précédent, qui l'enserre à l'Est et au Sud-Est, dans la région du haut Ouellé. Sa limite méridionale part de l'Itouri, puis suit à peu près la ligne de partage des eaux entre l'Ouellé-Oubangui et l'Itouri-Congo, pour atteindre ensuite l'Oubangui et le traverser près d'Imfondo ou Desbordesville, d'où elle remonte vers le Nord-Ouest jusque non loin de la haute Kadeï, où elle quitte la frontière nord du groupe bantou pour suivre la frontière orientale du groupe nigéro-camerounien jusque vers la haute Mambéré. Là commence la limite septentrionale du groupe oubanguien, qui marche à peu près droit vers l'Est jusque dans les vallées du Bahr-el-Ghazal, où elle s'arrête à la rencontre des groupes kordofanien et nilo-abyssinien. De

bantou; il est possible que les tribus connues sous ces deux noms parlent à la fois une langue bantou importée du Sud et une langue nilo-congolaise trouvée sur place.

nombreuses peuplades nègres remplissent ce vaste domaine, dont quelques-unes, comme les Zandé, possèdent une organisation relativement développée, mais dont la plupart sont encore à un stade très arriéré.

Toutes les langues du groupe possèdent des classes nominales à préfixes, analogues à celles du groupe nilo-congolais, quoique moins nettes. Elles sont attestées par la persistance plus ou moins générale de préfixes nominaux qui, cependant, ont souvent disparu, surtout au singulier. On ne paraît plus faire usage de pronoms distincts pour représenter les noms des différentes classes, sauf en zandé, mais on rencontre souvent, entre le nom et son complément ou son qualificatif, des particules qui semblent bien être des vestiges d'anciens pronoms de classe faisant aujourd'hui office d'éléments d'annexion ou de relation. Enfin le pronom de la 3e personne s'emploie comme déterminatif de la même façon que le pronom de classe dans les langues à système de classes intégral (banda, sango, zandé, mittou), ou bien l'on a un déterminatif qui rappelle la particule d'annexion (mandjia, baya).

Le nom ne se distingue essentiellement du verbe que là où le préfixe de classe nominale a persisté. Dans le cas contraire, qui est le plus fréquent dans beaucoup de langues (mandjia, baya, sango, banziri, bondjo, monjombo, séré, moungou), la distinction n'est assurée que par la place respective du nom et du verbe.

En général, l'injonctif est constitué par le simple radical verbal ou par un affixe préfixé au pronom sujet, comme en zandé ; l'aoriste se présente sous cette même forme ou prend un suffixe, parfois le préfixe *a* (sango, zandé) ; le parfait est toujours marqué par un suffixe ; certains aspects secondaires du verbe s'obtiennent par redoublement du radical (zandé, mandjia). La négation s'indique à l'aide d'une particule suffixée à la proposition.

Ordre des mots : le sujet précède le verbe ; le complément d'un verbe suit ce verbe ; le complément d'un nom suit ce nom, soit directement, soit précédé de la particule dont il a été question

plus haut ; parfois, en zandé, le nom est précédé de son complément, qui est précédé lui-même de la particule ; le qualificatif précède le nom qualifié si l'on entend laisser à celui-ci une valeur indéterminée (banda *i-mbin mbrata* « blanc cheval, un cheval blanc ») et le suit, avec ou sans particule de relation, dans le cas contraire (banda *mbrata di mbin* ou *mbrata se di-mbin* « le cheval blanc dont il est question ») ; le déterminatif suit toujours le nom déterminé, sauf en zandé ; s'il y a aussi qualificatif, le déterminatif se place généralement entre le nom et le qualificatif (banda, mandjia, baya, sango, etc.) ou exceptionnellement avant le qualificatif (zandé) ; le nom de nombre suit le nom de la chose nombrée.

Il semble exister des tons à valeur étymologique, c'est-à-dire qu'une même syllabe, émise sur des tons différents, peut représenter des idées distinctes, mais il ne paraît pas y avoir de tons musicaux à valeur grammaticale.

Il n'y a pas de désinences proprement consonantiques. Lorsque la racine se termine par une consonne, on lui suffixe, pour en faire un mot, une voyelle qui est en général semblable à la voyelle radicale.

En principe, le pluriel est indiqué dans les noms par le changement de préfixe ou par l'addition d'un préfixe aux mots qui n'en ont pas au singulier (banda : *mbrata* « cheval », plur. *a-mbrata* ; *bānda* « un Banda », plur. *a-bānda* ; zandé : *bọrọ* « être humain », plur. *a-bọrọ* ; *zānde* « un Zandé », plur. *a-zānde*). Mais, très fréquemment, le nombre n'est marqué que par le déterminatif, en sorte que rien ne peut alors faire distinguer le pluriel du singulier dans les noms indéterminés. D'après Struck, le gbaya formerait le pluriel des noms en changeant en *a* la voyelle terminale du singulier.

Voici les formes essentielles des pronoms dans les principales langues :

	Banda	Baya et Mandjia	Sango	Banziri	Bondjo	Zandé	Mittou
1re pers. sing.	*mo*	*mi, mô*	*mbi*	*ma*	*ma*	*mi*	*mu*
— plur.	*a*	*le*	*a*	*ya, a*	*a*	*a*	*ẹ*

2e pers. sing.	*bo, ẓe*	*me, ma*	*mo*	*mo, lo*	*mo, lo*	*mọ*	*mo*
— plur.	*e*	*ne*	*i*	*yi, i*	*i*	*o*	*i*
3e pers. sing.	*se, i*	*a, ya*	*lo*	*e*	*ye, e*	*kọ, li, u, sẹ*	*kọ, ọ*
— plur.	*nẓi*	*wa*	*a-la*	*wo, o*	*yo*	*yo, i, si*	?

Les renseignements que nous possédons nous permettent d'attribuer, avec les réserves nécessaires, 25 langues au groupe oubanguien :

1° le **mittou** (*mittu*) ou *abukaya* ou *avukaya*, dit parfois *mittu-madi* ou *amadi* (ne pas confondre avec le « madi » du groupe nilo-congolais ni avec le « morou » du même groupe parlé à Amadi-sur-Yeï), comprenant les dialectes des *Abukaya*, entre les Morou d'Amadi-sur-Yeï au Nord et les Abaka au Sud, des *Abaka*, entre les Bari à l'Est et les Zandé à l'Ouest, des *Luba*, entre les Abaka au Nord et les Logo au Sud, et enfin des *Amadi* au Aboukaya de l'Ouest, dans le district d'Amadi-sur-Ouellé entre 25° et 24° de longitude est de Paris (1864 Marno) ;

2° le **moungou** (*muñgu*) ou *maygo-muñgu* ou *moygo-muñgu* ou *mayogo*, parlé d'une part entre la haute Yeï et le haut Doungou par 4° latitude nord et 28° longitude est de Paris (dialecte *mundu*) et d'autre part au Sud des Mangbètou du centre par 3° lat. nord et 26° long. est de Paris (dialecte *mayogo*) (1864 Marno)[1] ;

3° le **gbaya** ou *beya*, dans le Dar-Fertit à l'Ouest de Dem-Ziber, sur le haut Chinko (1918 Struck) ;

4° le **kredj** (*kredẓ*) ou *kred'* ou *kreš* ou *kreyš* ou *fertīt*, dans le Dar-Fertit au Sud de Hofrat-en-Nahas et au Nord-Ouest du précédent, entre 9° et 8° lat. nord et 23° et 21° long. est de Paris : dialectes *ambãgo* ou *yãgbãgo*, *adẓa* ou *veya*, *ndoggo* ou *nduggo*, etc. (1829 Rüppell) ;

5° le **gobou** (*gobu*) ou *ñgobu* ou *agobbu*, sur le moyen Chinko au sud du gbaya (1888 Junker) ;

1. Le Père Vanden Plas prétend que les Moundou et les Mayogo parleraient le bangba, qui est une langue nilo-congolaise ; cette affirmation ne concorde pas avec les documents que nous possédons sur les dialectes moundou et mayogo, à moins que les Moundou et les Mayogo ne parlent le bangba en plus de leur langue maternelle, ce qui est bien possible.

6° le **gôlo** (*gōlo*) ou *gōro*, au Nord des Zandé de Zémio, par 6° lat. nord (1829 Rüppell);

7° le **séré** (*sere*) ou *šere* ou *basiri* ou *basili*, parlé dans quelques villages répandus parmi les Gôlo et au Sud-Est de leur pays (1912 Westermann, qui donne du séré sous le nom de gôlo);

8° le **ndakko** ou *andakko*, parlé au voisinage du séré et très voisin de celui-ci (1888 Junker);

9° le **barambo** ou *baramo* ou *balẹmbo* ou *abarambo*, parlé dans quelques villages au voisinage des Zandé, notamment à l'Ouest de Tamboura et sur la rive sud de l'Ouellé par 25° long. est de Paris (1888 Junker);

10° le **zandé** (*zãnde*) ou *azãnde* ou *n'amn'am* ou *man'an'a* ou *omad'aka* ou *babuñgera* ou *makkarakka* ou *makraka* ou *makarka* ou *makalaka* ou *digga*, parlé par une population de plus de 200.000 individus, répartie entre de nombreuses tribus conquérantes ou vassales dont les principales sont les *Avuñgura* ou *Avoñgara*, les *Aband'a* ou *Abanza* et les *Ad'o*, habitant sur la rive septentrionale du Mbomou à partir de Rafaï et en amont et sur la zone comprise entre le Mbomou et l'Ouellé du 27° au 22° long. est de Paris (1861 Petherick);

11° le **dendi**, sur le Mbomou entre Rafaï et Bangasso (néant);

12° le **sakara** ou *nsakkara*, au Nord du précédent et à l'Ouest dans la région de Bangasso et des Patri (1895 Comte);

13° sous réserves : le **sabanga** (*sabãga*) ou *manza* ou *mãñga* ou *bamãñga*, parlé d'une part au milieu des Banda par 7° lat. nord et 19° long. est de Paris, d'autre part au coude de l'Oubangui (région de Fort-de-Possel) et enfin près et en aval de Bangui (néant);

14° le **sango** (*sãñgo*), dialectes *buraka*, *yakoma*, *bãgi*, etc., parlé sur les deux rives du Mbomou près de son confluent avec l'Ouellé et, en aval de ce confluent, sur les deux rives de l'Oubangui (régions de Ouango, Mobaye, Banzyville), et répandu comme langue commerciale tout le long de l'Oubangui jusque vers Imfondo ou Desbordesville (1906 Demombynes d'après Decorse);

15° le **gbandi** (*gbãndi*) ou *ñgbãndi* ou *moñgbãndi* ou *mo-*

ṅgwãndi, au Sud du précédent, jusque vers le 3° lat. nord, entre les 21° et 18° long. est de Paris (1921 Tanghe);

16° le **banda** (*bãnḍa*) ou *abãnda*, parlé au Nord du Mbomou-Oubangui, dont il est séparé par le zandé, le sakara, le sango, le maka, le banziri, et dont il atteint les rives en quelques endroits près du coude de Fort-de-Possel et de Bangui, s'étendant en longueur du 22° au 17° environ de long. est de Paris et dépassant au Nord le 8° lat. nord; population de 200.000 individus environ, répartie entre de nombreuses tribus (Bodo, Biri, Vidri, Ouassa, Togbo, Mbriya, Linda, Yakpa, Langba, Goundjou, Pagoua, Ndopa, Bendi, Boubou ou Bougou, Mbala, Ngao, Marouba, Dakoa, Dakpa, Ngapou, Gba, Langouassi, Djemmi, Tané, Tombaggo, Mbagga ou Gbaga, Ngola ou Ngoura, Mbi, Ka, Ndi, Mbrou ou Bourou, Mbré ou Mbélé, Moria, Mbada ou Pata ou Ouadda, Ndri, Nguéré ou Nguélé ou Bouzérou, etc.) (1902 Truffert);

17° sous réserves : le **maka** ou *bamaka* ou *saka*, parlé sur l'Oubangui par une petite tribu fixée entre les Sango et les Banziri (néant);

18° le **banziri** (*bãnẓiri*) ou *gbãnẓiri*, sur l'Oubangui en aval de Mobaye et des Maka jusque près et en amont de Fort-de-Possel et, plus en aval, à hauteur et en aval du confluent de la Lobaye (1906 Demombynes d'après Decorse);

19° le **bondjo** (*bõdẓo*) ou *bõdẓyo* ou *mbwaga* ou *gmbwaga* ou *mbwaka* ou *mpagga*, sur la basse Lobaye à l'Ouest des Banziri du Sud (1911 Ouzilleau, Calloc'h);

20° le **monjombo** (*monẓombo*) ou *mondẓombo* ou *modẓambo*, dit aussi « bondjo » comme le précédent, parlé le long de l'Oubangui et à l'Ouest, de Bétou ou Monjimbo jusqu'à Imfondo ou Desbordesville inclus (1911 Cottes et Poutrin, Ouzilleau, Calloc'h);

21° le **mandjia** (*mãdẓya*) ou *mãdẓa* ou *gbea*, parlé entre les 17° et 15° long. est de Paris, du 7° au 4° lat. nord, à cheval sur le bassin du Chari et celui de l'Oubangui, par de nombreuses tribus réparties en quatre sections : Baza ou Baja, Bakka, Bidigri et Mombé (1908 Giraud);

22° sous réserves : le **yanguéré** (*yàṅgere*) ou *bosiṅgene* ou *kuruma*, à l'Ouest du précédent (néant);

23° le **baya**, parlé à l'Ouest du précédent, entre les 14° et 12° long. est de Paris, et dans deux enclaves méridionales, au milieu de parlers bantou, chez les Bangandou (entre la Ngoko et la Sanga) et les Ikassa de la haute Likouala (1896 Clozel);

24° le **goundi** (*gundi*), au Sud du yanguéré, sur la haute Lobaye (région de Bambio) et sur la haute Sanga (région de Bayanga), au milieu de parlers bantou (1911 Ouzilleau);

25° le **bomassa** (*bomasa*), sur la moyenne Sanga à l'Est de Ouesso, au milieu de parlers bantou (1911 Ouzilleau).

BIBLIOGRAPHIE. — Il n'existe pas d'ouvrage d'ensemble sur les langues oubanguiennes. Quelques-unes ont fait l'objet de bonnes études spéciales, notamment : le zandé (LAGAE et VANDEN PLAS, *La langue des Azandé*, Gand, 1921-1922, 2 vol. in-8) ; le sango (CALLOC'H, *Vocabulaire français-sango et sango-français précédé d'un abrégé grammatical*, Paris, 1911, in-8) ; le banda (COTEL, *Dictionnaire français-banda et banda-français précédé d'un essai de grammaire banda*, Brazzaville, 1907, in-18) ; le banziri, le bondjo et le monjombo ensemble (CALLOC'H, *Vocabulaire français-gmbwaga-gbanziri-monjombo, précédé d'éléments de grammaire*, Paris, 1911, in-8) ; le mandjia (même auteur, *Vocabulaire français-gbéa précédé d'éléments de grammaire*, Paris, 1911, in-8). A citer aussi : EBOUÉ, *Langues sango, banda, baya, mandjia*, Paris, 1918, in-16 obl.

VII. — *Groupe chari-ouadaïen* (12 langues).

Le domaine de ce groupe s'étend, d'une façon générale, au Sud de la partie occidentale du groupe nilo-tchadien et au Nord de la partie centrale et occidentale du groupe oubanguien. A ses extrémités est et ouest, il est interpénétré par le domaine de ce dernier groupe et par celui du groupe nigéro-tchadien ; il est entamé de plus par une enclave que constitue le groupe charien· Son extension territoriale est peu considérable et se trouve diminuée encore par la présence de quelques îlots de langue peule (du groupe sénégalo-guinéen), notamment au Baguirmi, et de langue arabe. Les populations parlant des langues chari-ouadaïennes sont composées uniquement de nègres dont la plupart

sont fort arriérés, mais dont quelques-uns, comme les Baguirmiens, ont atteint un degré de civilisation relativement avancé. Leur nombre total est assez élevé ; à eux seuls, les Baguirmiens atteignent ou dépassent le chiffre de 100.000, et les Sara sont vraisemblablement au moins aussi nombreux.

Les langues chari-ouadaïennes constituent l'un des très rares groupes négro-africains dans lesquels on ne trouve pas actuellement de classes nominales ni de pronoms de classe. Toutefois l'existence ancienne de classes paraît attestée par le fait que certains dissyllabes, jouant aujourd'hui le rôle de radicaux, comportent visiblement un préfixe n'appartenant pas à la racine, lequel préfixe varie selon les langues ou les dialectes et revêt souvent une forme que l'on rencontre devant des noms de même catégorie, comme préfixe de classe, dans des groupes à classes nominales. C'est ainsi que *ma* dans le mot sara, baguirmien et goula *ma-nda* « jeune fille » et *mo* dans le baguirmien *mo-belo* « vieille femme » rappellent le préfixe de la classe humaine du bantou et d'autres groupes, tandis que *nda* et *bel* sont assurément les racines, comme il résulte d'une comparaison avec d'autres langues. A signaler de même : sara *ndo* et *gi-ndo* « nuit » (en horo *lu-ndo*) ; *na-he* « lune » (en ndouka *nze-he*, en bongo *ni-hi*, en babalia *ne-fe*, en kenga *na-fa*, en lis *na-fe*, en baguirmien *na-pi*) ; *yo-ñgọ* et *nu-ñgọ* « os » (kaba et ndouka *nu-ñgọ*, goula *n'o-ñgọ*, lis *to-ñgọ*, baguirmien *tu-ñọ* et *t'u-ñgọ*, horo *ñgọ*) ; *ñga* « terre » (ndouka *ñga*, goula *na-ñga*, baguirmien *na-ñge*), etc.

Le nom n'est distingué du verbe, d'une façon essentielle, que par la place respective qu'ils occupent dans la phrase.

La conjugaison du verbe présente une certaine originalité. Il en existe divers types selon que la racine verbale est du type « consonne + voyelle + consonne », du type « consonne + voyelle » ou du type « voyelle + consonne » ou « voyelle ». Dans le premier cas et le troisième, on suffixe une voyelle à la racine ; dans le second, le troisième et le quatrième, on préfixe un élément *ki* ou *ku* (*k* devant une voyelle), qui disparaît au parfait ; dans le premier cas, c'est la forme du pronom sujet qui

permet de distinguer l'aoriste du parfait. Il convient de noter à ce propos que la voyelle du pronom sujet se modifie ou même disparaît selon les aspects verbaux, que le pronom sujet de la 1[re] personne est identique au pluriel à celui de la 3[e] personne (la distinction étant faite éventuellement à l'aide d'un suffixe *ki* ou *gi* ajouté au verbe à la 1[re] personne) et qu'à la 2[e] personne, le pronom sujet (qui parfois ne s'exprime pas au parfait) est généralement le même pour les deux nombres (la distinction étant faite par l'addition au verbe du suffixe *ki* ou *gi* au pluriel). Il arrive souvent que les aspects principaux du verbe ne se différencient que par la forme du pronom sujet ; à la 3[e] personne, si le sujet est un nom, aucune distinction n'est possible à la première conjugaison. D'autre part, il existe un présent à préfixe (baguirmien *el*, sara *tu*) et des aspects secondaires ou négatifs à suffixes ; ces suffixes sont parfois séparés du verbe par son complément.

Ordre des mots : le sujet précède le verbe ; le complément d'un verbe suit ce verbe ; le complément d'un nom suit ce nom, avec, éventuellement, préfixation (baguirmien *an*) ou suffixation (bongo *gi*), au nom ou pronom complément, d'une particule précisant l'annexion ; le qualificatif suit en général le nom qualifié, mais peut aussi le précéder, et cela dans la même langue ; le déterminatif et le nom de nombre suivent le nom déterminé ou nombré.

Les désinences consonantiques ne sont pas rares ; mais, en général, les racines se terminant par une consonne ne deviennent des mots qu'après suffixation d'une voyelle qui, souvent, est identique à la voyelle radicale ou à la voyelle de l'ancien préfixe de classe et ne semble avoir alors que la valeur d'une voyelle d'appui, mais qui, d'autres fois, est différente de la voyelle radicale et paraît constituer alors un suffixe de dérivation (ainsi baguirmien *dona* « force physique » et *dono* « influence morale ou politique » ; *tuma* « marcher » et *tumo* « marcher à la dérobée »). Quelle que soit sa valeur, cette voyelle tombe devant un complément, un qualificatif ou un déterminatif, si la voyelle radicale est accentuée, et subsiste dans le cas contraire : baguirmien *dóna*

et *dóno* donnent *dón ñgólo* « grande force » ou « grande influence » ; *kága* « arbre » donne *kág muta* « trois arbres », mais *kaga* « panthère » donne *kaga muta* « trois panthères ».

Le pluriel des noms se marque par un suffixe spécial (*-ge* en baguirmien et sara), qui entraîne la chute de la voyelle terminale non radicale si la voyelle radicale est accentuée (baguirmien *bísi* « chien », plur. *bís-ge* ; *sinda* « cheval », plur. *sinda-ge*).

Les pronoms personnels, dans leur forme essentielle, sont :

	Bongo	Baguirmien	Ndouka	Kaba	Lis	Sara
1re pers. sing.	*ma*	*ma*	*ma*	*ma*	*ma*	*ma*
— plur.	*ge*	*d'e*	*dze*	*dže*	*d'e*	*ze*
2e pers. sing.	*i*	*i*	*i*	*i*	*yi*	*yi*
— plur.	*hę*	*se*	*se*	*se*	*ze*	*se*
3e pers. sing.	*ba*	*ne*	*ne*	*ne*	*ne*	*ne*
— plur.	*ye*	*d'e*	*dzi*	*de*	*de*	*zi*

Le groupe chari-ouadaïen paraît renfermer 12 langues :

1° le **goula** (*gula*) ou *bolgo* ou *disa* (cette dernière appellation commune au goula et au lis), parlé d'une part au Nord-Ouest du kredj et d'autre part sur la rive droite du Bahr-Salamat au Sud du saba (1906 Demombynes d'après Decorse) ;

2° le **bongo** (*bòñgo*) ou *abõñgo* ou *oboñg* ou *akuma* ou *dōr*, dans la région montagneuse d'où sort à l'Est le Bahr-el-Arab (1857 Von Heuglin) ;

3° le **ndouka** (*nduka*) ou *ndukwa* ou *ndokwa* ou *ndokoa*, sur la rive droite du haut Bamingui (1906 Demombynes d'après Decorse) ;

4° sous réserves : le **saba** ou *sabañg* ou *bañgbay* ou *d'oñgor*, vallée du Bahr-Salamat (1857 Barth) ;

5° et 6° le **kaba** ou *kabba*, sur le Bahr-Salamat, près de son confluent avec le Chari, et le **horo**, rive droite du Chari en amont de Fort-Archambault (1906 Demombynes d'après Decorse) ;

7° le **sara**, entre le haut Chari et le haut Logone par 9° environ de lat. nord et dans une enclave située à l'Est du moyen Chari entre les Kaba et les Goula du Bahr-Salamat (1898 Delafosse) ;

8° le **ngama** (*ṅgama*) ou *dagba*, au Sud-Est des Sara, entre le Bahr-Sara et le Gribingui (1906 Demombynes d'après Decorse et Bruel) ;

9° le **lis** ou *lisa* ou *disa* ou *tār-lis* (langue des Lis), dialectes *medogo*, *kuka* et *bulala*, parlé sur la basse Batha, le lac Fitri et au Sud (1906 Demombynes d'après Decorse) ;

10° le **kenga** (*keṅga*), entre la Batha et le Chari (1909 Gaden) ;

11° le **barma** (*bārma*) ou « baguirmien » ou *bagarmi* ou *begarmi* ou *bagirmi* ou *bagrimma* ou *bögre* ou *tomo*, dans le Baguirmi, à l'Est du bas Chari (1817 Vater d'après Seetzen);

12° le **babalia** ou *mbrak*, sur la rive droite du Chari, en aval du précédent, entre Fort-de-Cointet et Fort-Lamy (1906 Demombynes d'après Decorse).

BIBLIOGRAPHIE. — Aucun ouvrage d'ensemble sur le groupe, mais des vocabulaires de plusieurs langues, recueillis par le Dr Decorse, figurent, accompagnés de notes grammaticales, dans GAUDEFROY-DEMOMBYNES, *Documents sur les langues de l'Oubangui-Chari* (II, Groupe barma), Paris, 1906, in-8. Il existe de plus de DELAFOSSE un *Essai sur le peuple et la langue sara*, Paris, 1898, in-8, et de GADEN un très bon *Essai de grammaire de la langue baguirmienne suivi de textes et de vocabulaires*, Paris, 1909, in-8.

VIII. — *Groupe charien* (15 langues).

Très restreint comme superficie, ce groupe forme le long du moyen et du haut Chari une enclave entre les langues nord-occidentales et sud-occidentales du groupe précédent, enclave se soudant sur le moyen Logone au groupe nigéro-tchadien. Il comprend de plus un îlot isolé à Afadé, au Sud et près du lac Tchad, dans le triangle Goulfeï-Kousseri-Dikoa. Les langues dont il se compose, fort mal connues encore, sont parlées par de petites tribus nègres en général très arriérées.

Il semble bien exister des classes de noms caractérisées par des préfixes et, dans plusieurs langues du groupe tout au moins, par des pronoms de classe spéciaux (saroua, toumok, somraï et ndam *n-am*, mana *r-im*, afadé *am* et *l-ẹm*, gabéri *k-am*, tounia et

niellim *n'-um* « eau » ; saroua *a-dwa*, toumok et ndam *a-ru*, somraï *be-da*, gabéri *ka-la* « bois » ; boua *m-bel*, tounia *li-bili* « chemin », etc.).

Le nom se distingue du verbe en général à l'aide des préfixes nominaux de classe et aussi par la place respective de l'un et de l'autre.

Les documents existants ne permettent pas de discerner le mode de conjugaison.

Ordre des mots : le nom ou pronom sujet précède le verbe ; le nom ou pronom complément d'un verbe suit ce verbe ; le nom ou pronom complément d'un nom se place avant ce nom, comme dans le groupe nilo-congolais et, en partie, dans le groupe nilo-tchadien, et contrairement à ce qui a lieu dans les cinq autres groupes étudiés jusqu'ici (saroua : *kwo-lẹ* « maison », *m-bu* « bouche », *kwo-lẹ m-bu* « bouche de maison, porte » ; boua : *lu* « maison », *m-mu* « bouche », *lu m-mu* « porte » ; mana *to m-mo*, niellim *li m-mu*, somraï *ku-lu bu*, « porte ») ; le qualificatif suit le nom qualifié, soit directement, soit précédé lui-même d'une particule qui est peut-être un indice de classe ; le déterminatif et le nom de nombre suivent le nom déterminé ou nombré.

Les désinences consonantiques sont assez rares ; la plupart du temps — mais non toujours — elles semblent ne se rencontrer que si la consonne finale est une sonante.

Le pluriel des noms paraît s'obtenir, tantôt par changement du préfixe de classe, tantôt, et notamment pour les noms dépourvus de préfixe, par l'addition d'un suffixe de pluralité ou par modification de la désinence vocalique.

Les documents existants ne permettent pas de reconnaître la forme exacte des pronoms.

Le groupe semble renfermer 15 langues, dont 5 parlées sur la rive droite ou à l'Est du Chari, 5 sur la rive gauche, 4 plus à l'Ouest entre Chari et Logone et 1 à l'Ouest et près de l'embouchure du Chari :

1° à 5° le **sokoro** ou *bedaṅga*, à l'Est du Baguirmi, entre Melfi et Bédanga inclus (1912 Benton d'après Barth) ; — le **saroua**

(*sarwa*) ou *sẹrwa* ou *kara*, en amont du babalia, de Fort-de-Cointet en aval à Boulaye en amont (1906 Demombynes d'après Decorse); — le **fania** (*fan'a*) ou *fan'ã* ou *koke*, à l'Est de Melfi (néant); — le **boua** (*bwa*) ou *boa*, région de Korbol, de Boulaye en aval jusqu'au confluent du Bahr-Salamat en amont (1905 Bruel) ; — le **mana**, près de Fort-Archambault (1906 Demombynes d'après Decorse) ;

6° à 10° le **miltou** (*miltu*) ou *kuñ*, de Fort-de-Cointet en aval jusqu'au 10° lat. nord environ, et le **niellim** (*n'ellim*) ou *n'illem*, en amont du précédent jusqu'à Fort-Archambault (1905 Bruel) ; — le **tounia** (*tun'a*), en amont du précédent, et, sous réserves, le **gori**, de situation géographique mal déterminée (1906 Demombynes d'après Decorse) ; — sous réserves : l'**arétou** (*aretu*) ou *rotu* ou *leto*, en amont du tounia (néant) ;

11° à 14° le **ndam** ou *dam*, au Sud du miltou et à l'Ouest du niellim ; le **somraï** (*somray*) ou *somrẹ*, à l'Ouest du précédent ; le **gabéri** (*gaberi*) ou *tšire* ou *t'ẽre* ou *masa* (cette dernière appellation partagée avec le mousgou, du groupe nigéro-tchadien), au Sud du somraï ; le **toumok** (*tumok*) ou *tummok* ou *tumak* ou *maga*, entre le pays des Gabéri et le Bahr-Sara (pour ces quatre langues, 1905 Bruel) ;

15° l'**afadé** (*afade*) ou *affade*, dans le canton de même nom, entre le bras du bas Chari dit Serbevel à l'Est et la Yadséram ou rivière de Dikoa à l'Ouest, près de la rive méridionale du Tchad (1808 Seetzen).

BIBLIOGRAPHIE. — Rien n'a été publié sur les langues de ce groupe, en dehors des vocabulaires, en général très fragmentaires, contenus dans : U. J. SEETZEN, *Monatliche Korrespondenz*, Band XXII, 1808 (pp. 269-275) ; J. S. VATER (d'après le précédent), *Mithridates*, Leipzig, 1817 ; G. BRUEL (numérations), *Le Cercle du Moyen-Logone*, Paris, 1905, in-12 ; GAUDEFROY-DEMOMBYNES (d'après Bruel et Decorse), *Documents sur les langues de l'Oubangui-Chari*, Paris, 1906, in-8 (III et surtout VII et VIII) ; BENTON (d'après des vocabulaires inédits recueillis par Barth en 1852), *Notes on some languages of the Western Sudan*, Oxford, 1912, in-18 (Part III).

IX. — *Groupe nigéro-tchadien* (31 langues).

Ce groupe fait suite à l'Ouest aux groupes nilo-tchadien, chari-ouadaïen et charien ; il pousse, au Sud, une pointe jusque dans le Sud de l'Adamaoua avec le batta et contourne à l'Est et au Nord le groupe nigéro-camerounien, auquel il semble abandonner les pays bas et les vallées pour se réserver les montagnes et les plateaux, — observation qui ne s'applique pas aux langues nord-orientales du groupe nigéro-tchadien (mousgou, klessem et kotoko du bas Logone, yédina et koûri du Tchad). Dans la direction de l'Ouest, avec le haoussa, il atteint le Niger sur une petite partie de son cours, des deux côtés de l'embouchure de la rivière de Sokoto, pour s'en éloigner aussitôt et côtoyer, dans la direction du Sahara, le groupe nigéro-sénégalais, représenté là par le zerma et le songoï. Au Nord, la limite avec le berbère est assez indécise, plusieurs fractions de Touareg parlant également le haoussa et le tamacheq ou même, comme beaucoup de Kel-Oui, ne faisant guère usage que de la première de ces langues.

Bien que l'étendue territoriale de ce groupe soit moins vaste que celle de certains autres, il présente une très grande importance, en ce sens que la densité de la population est relativement fort élevée dans la majeure partie de son domaine propre, également par suite du grand nombre de ses parlers, enfin du fait que l'un de ceux-ci, le haoussa, idiome maternel d'une société en partie très évoluée, est devenu la principale langue d'extension et de civilisation du Soudan central. Par contre, les autres langues du groupe, dont plusieurs paraissent être en voie de disparition, sont parlées en général par des populations très frustes et souvent barbares.

Tous les peuples qui ont comme parler national l'une des langues du groupe, y compris celui des Haoussa, appartiennent incontestablement à la race nègre. Les idiomes en usage dans les montagnes du Baoutchi et de l'Adamaoua semblent être les représentants les plus purs de l'ancien fonds commun. Au con-

traire, la langue haoussa a bénéficié de l'évolution du peuple haoussa, évolution à laquelle le contact des races de l'Afrique du Nord n'a probablement pas été étranger. On a pensé souvent que ce même contact avait pu introduire dans la langue des modifications dues à une influence sémitique ou libyco-berbère, comme l'emploi d'un genre féminin à désinence *a* et d'un pronom *ta* spécial à ce genre, quelques autres pronoms, l'usage d'une particule d'annexion *n*, etc. Toutefois, une étude attentive des autres langues du groupe, et en particulier de celles qui sont demeurées le plus à l'écart des contacts chamito-sémitiques, montre que ces phénomènes ne sont point spéciaux au haoussa. Lorsque nous connaîtrons mieux l'angas, le bouta, le bola, le batta et d'autres langues nigéro-tchadiennes dont nous ne savons guère aujourd'hui que le nom, il deviendra loisible de trancher la question, souvent agitée mais impossible à résoudre actuellement de façon sûre, de l'influence des parlers chamito-sémitiques sur le haoussa. En tout cas, même si cette influence vient à être démontrée quelque jour, il paraît certain dès maintenant qu'elle n'a pas été suffisante pour que l'on puisse distraire le haoussa de l'ensemble des langues négro-africaines et, en particulier, du groupe nigéro-tchadien.

Il convient de noter que le domaine de ce groupe renferme des îlots importants de langue peule (groupe sénégalo-guinéen), notamment dans l'Adamaoua et, à un degré moindre, dans les provinces de Gober et de Sokoto.

L'existence de classes nominales à préfixes se reconnaît facilement dans toutes les langues du groupe, mais elle s'y présente, selon les langues, à des stades très divers d'évolution. Souvent, les pronoms de classe ont persisté, même là où les préfixes de classe ont disparu ou tendent à disparaître, et ils sont employés comme sujets et régimes aussi bien que comme déterminatifs et comme préfixes d'accord de l'adjectif ; le cas est attesté particulièrement en yédina, en mousgou, et, à un degré moindre, en kotoko, en mandara, en margui, en batta, etc. ; il l'est également en haoussa, mais, dans cette langue, comme en d'autres

d'ailleurs et notamment en mousgou et en bola, le nombre des classes est réduit à deux pour le singulier, avec une classe commune au pluriel, au moins du point de vue grammatical, et l'on n'a plus au singulier que deux pronoms de classe, dont l'un (*ta* en haoussa, *t* en mousgou et bola) représente les noms des êtres féminins et d'un certain nombre d'objets sans sexe (généralement à désinence *a* en haoussa, *i* en mousgou), ainsi que les noms propres de localités, l'autre pronom (en haoussa *ši*, *sa*, *ya* ou *na* selon les cas, en mousgou *ni*, *na* ou *a*, en bola *se* ou *ni*) représentant tous les autres noms (y compris, en haoussa, un nombre appréciable de noms terminés par *a* et, en particulier, de noms abstraits dérivés en *ta*). Même dans les langues comme le haoussa et le bola, par exemple, qui n'ont plus que deux classes grammaticales, pouvant à la rigueur recevoir l'étiquette, consacrée chez nous, de genres masculin et féminin, il y a lieu de noter la présence de préfixes, se rapportant à d'autres catégories plus anciennes, qui, ayant perdu aujourd'hui leur signification première, semblent à première vue faire partie du radical. On dit en haoussa *mu-tum na-gari* « une personne bonne », *me-ži na-gari* « un homme bon », *ma-tše ta-gari* « une femme bonne », mais, dans la première syllabe de *mu-tum*, *me-ži*, *ma-tše*, il est aisé de reconnaître le préfixe de la classe humaine si fréquent dans l'ensemble des langues négro-africaines (comparez haoussa *mu-tum* « être humain », bola *mu-ndu* « femme », margui *ma-ndu* et zani *mi-nde* « être humain », et diverses langues bantou *mu-ntu* « être humain » ; haoussa et batta *ma-tše*, mandara *mu-ksa* et kongo *ṅ-ke* « femme »). Ce même préfixe d'une ancienne classe humaine a continué à trouver son emploi comme préfixe des noms d'agent ou de détenteur, sous la forme *ma* ou *may* en haoussa, *ma* en mandara, *mo* ou *mu* en mousgou, *me* ou *mi* en kotoko, *moy* ou *am* en bola, etc.

Le nom, lorsque le préfixe de classe a disparu ou a été, en quelque sorte, incorporé au radical, ne se distingue essentiellement du verbe que par la place respective de l'un et de l'autre. Parfois il est fait usage de désinences proprement nominales et, moins rarement, d'affixes verbaux de conjugaison.

La conjugaison présente peu d'unité. Toutefois il est un caractère commun à tout le groupe : c'est que l'aspect verbal le plus employé, l'aoriste, se compose du simple énoncé du radical verbal, sans affixe, et sert aussi d'injonctif. Les autres aspects sont marqués, tantôt par des préfixes, tantôt par des suffixes, tantôt par l'emploi de telle forme de pronom sujet de préférence à telle autre, tantôt par le premier et le troisième procédés réunis. La négation s'obtient en général soit en suffixant *à la proposition* une particule négative, soit en préfixant *au sujet* une particule qu'on répète parfois à la fin de la phrase ; quelquefois on substitue, devant le radical verbal, un préfixe négatif à un préfixe affirmatif (batta). Parfois le pluriel se distingue du singulier dans les verbes, soit par l'addition d'un suffixe nasal spécial (bola à toutes les personnes, yédina à la 2e), soit par le changement des voyelles radicale et terminale (mousgou). Parfois aussi on ajoute un suffixe au radical lorsque le verbe est employé sans complément direct (bola). Parfois on a des formes verbales caractérisées par le redoublement du radical (mousgou, mandara, haoussa). Parfois enfin le parfait se distingue de l'aoriste par le fait qu'au premier de ces aspects on substitue à la voyelle radicale, selon la personne, la voyelle essentielle du pronom (kotoko).

Ordre des mots : 1° le nom ou pronom sujet précède le verbe [1] ; lorsque le sujet est un nom, on exprime généralement après lui le pronom sujet, ce qui souvent est nécessaire pour préciser l'aspect du verbe, la forme du pronom sujet étant parfois la seule chose qui le précise : haoussa *kare ya nema* (chien il chercher)

1. Sauf, disent la plupart des auteurs, au passif. Mais, en réalité, le soi-disant passif du haoussa, du mandara, du bola, etc., est un verbe actif à sujet indéterminé *a*, qui se retrouve en yédina, en mousgou et en batta comme pronom commun à toutes les classes, en kotoko et en mandara comme pronom unifié, et qui correspond à notre « on » ; et le verbe a pour régime — et non pour sujet — ce qui serait le sujet de notre passif. Cela est si vrai que, si c'est un pronom qui subit l'action, il prend la forme du pronom complément : en haoussa *a nema kare* et *a nemaši* signifient exactement « on cherche le chien » et « on cherche lui » et non pas « le chien est cherché », « il est cherché ».

« le chien cherche », *kare ši nema* (chien lui chercher) « que le chien cherche » ; [1] — 2° le nom ou pronom complément d'un nom suit ce nom, soit directement, soit le plus souvent en intercalant entre eux une particule d'annexion qui est généralement *n*, mais qui peut aussi varier selon la classe du nom complété, surtout lorsqu'on veut insister sur l'idée de propriété ou de dépendance (yédina *ñgö* et *ru* ; kotoko *na* ou *en*, *e*, *a* et *ya* ; mousgou *ñga*, *na*, *ne*, *ta*, *la* ; mandara *na*, *a*, *da* ; haoussa *na* ou *n* et *ta* ou *l* ou *r*) ; cette particule n'est donc pas autre chose que l'une des formes du pronom ou indice de classe : haoussa *birni n sarki* « la ville du roi », *matše n sarki* « la femme du roi », où *n* n'est pas l'équivalent de *n* berbère marquant le génitif, mais de l'ancien pronom *na* de la 1re classe qu'on retrouve dans *sarki na-gari* « le roi bon » ; de même, on a *birni na sarki* « la propre ville du roi, la ville celle du roi », mais *matše ta sarki* ou *matše l sarki* « la propre femme du roi » ; — 4° le qualificatif suit le nom qualifié si l'on veut faire entendre que celui-ci est déterminé et le précède dans l'hypothèse contraire ; dans le premier cas, on fait accorder l'adjectif en classe avec le nom, soit en lui préfixant un indice de classe approprié, soit en lui donnant une désinence spéciale, soit par les deux procédés ensemble ; dans le second cas, l'adjectif reste en général invariable et c'est le nom que l'on fait précéder de l'indice de classe : haoussa *meži na-gari* « l'homme bon », *matše ta-gari* « la femme bonne » ; *gida sabo* « la maison neuve », *riga sabua* « la blouse neuve » ; *gari n-meži* « un homme bon », *gari l-matše* « une femme bonne » ; *sabo n-gida* « une maison neuve », *sabo l-riga* « une blouse neuve » ; — 5° le déterminatif suit le nom déterminé (yédina *köli* « chien », *köli ngö* « le chien » ; haoussa *kare* « chien », *kare n* « le chien ») ; — 6° le nom de nombre se construit de même et prend souvent un préfixe de classe, le nom de la chose nombrée pouvant rester au singulier : « trois femmes », yédina *ñgö-röm ka-könnö*, kotoko *ge-nem ga-xir*, mous-

1. A la 2e classe, toutefois, on aura la même forme (*karia ta nema*) pour « la chienne cherche » et « que la chienne cherche ».

gou *mu-ni mu-bu*, mandara *mu-ksa kadẓe*, bola *mu-ndu kunu*, batta *ma-tše ma-kin*, haoussa *ma-tše uku*.

On rencontre assez souvent des désinences consonantiques, mais le plus souvent, dans ce cas, la consonne finale est une sonante (liquide, nasale ou semi-voyelle) ou un *s* ; cependant en mousgou, en kotoko, en batta et en zani, on a des terminaisons *f*, *d*, *t*, *g*, *k*, etc. Le changement de nombre ou de classe dans les noms et les adjectifs ou d'acception dans les verbes peut amener une modification de la voyelle radicale ou d'une consonne radicale ou des deux à la fois, ou encore l'addition d'une voyelle entre deux consonnes radicales : yédina *fo* « village », plur. *fu-yu* ; *wule* « enfant », plur. *wayl-u* ; kotoko *ge-nem* « femme », plur. *ge-nam* ou *ga-nam* ; *böskwan* « cheval », plur. *basakwa* ; mousgou *sa* « boire », *ši* « être bu » ; *wẹl* « vieillard », *uli* « vieille femme » ; bola *došo* « cheval », plur. *dowi* ; haoussa *doki* « cheval », plur. *dawaki*.

Le pluriel se marque dans les noms soit par l'addition d'un suffixe de pluralité au singulier, soit par le changement de la désinence, soit par une modification de la racine (voir plus haut), soit par substitution d'un préfixe à un autre ou par chute du préfixe, soit par redoublement de la seconde consonne radicale (haoussa *tufa* « tissu », plur. *tufafi* ; *iri* « espèce », plur. *irari*), soit par plusieurs de ces procédés à la fois.

Voici la forme essentielle des pronoms dans quelques-unes des langues du groupe :

	Yédina	Koûri	Kotoko	Mousgou	Mandara	Gamergou	Bola	Haoussa
1re pers. sing.	*wu*	*u*	*wu, in*	*mu, ā*	*ye, wa*	*i, wa*	*i, ni, wo*	*i, ni, wa*
— plur.	*yę, nay*	*ne*	*mi*	*mi*	*ma*	?	*mu*	*mu*
2e pers. sing. 1re cl.	*gu*	*ku*	*ken, ku*	*ku*	*ka*	*ka*	*ka*	*ka*
— — 2e cl.	...	...	*ki*	...	*ki*	*ke*	...	*ki*
— plur.	*gu, woy*	*kay*	*kan, kun*	*ki*	*ku*	*ku*	*ku*	*ku*
3e pers. sing. 1re cl.	*ni*	*ni*	*ni*	*ni*	*na*	?	*ni, se*	*ši, sa, ya, na*
— — 2e cl.	*di*	*di*	*di*	*ta*	*da*	?	*ta*	*ta*
— commun	*a*	*a*	*a*	*a*	*a*	?	?	*a*
— plur.	*yi, dan*	*tan*	*ye, ti*	*e, di*	*te*	*tā*	*su*	*su*

On notera la présence fréquente — et non point seulement en haoussa — d'un pronom spécial pour la 2ᵉ pers. du sing. 2ᵉ classe, ou féminin.

L'état actuel de nos connaissances nous conduit à attribuer 31 langues au groupe nigéro-tchadien :

1° le **yédina** (*yedina*) ou *yeddina* ou *buduma* au *budduma* ou *bidduma*, parlé dans les îles du centre, du Nord et de l'Est du lac Tchad, ainsi que sur ses rives nord et nord-est (1854 Koelle) ;

2° le **koûri** (*kūri*) ou *kūra* ou *kāle* ou *kādi* ou *ṅgāne* ou *karka*, parlé dans les îles et sur les rives de la partie sud-est du lac Tchad (1906 Demombynes d'après Decorse) ;

3° le **kotoko** ou *logone* ou *logon* ou *malagom* ou *buso* ou *bedde* ou *bal* ou *makari* (dialectes *ṅgala* ou *sao* ou *sō*, *semsig* ou *kusri*, *gulfey*, etc.), sur les deux rives du bas Logone en aval de Mousgoum et sur les deux rives du bas Chari depuis le lac Tchad jusqu'à Fort-Lamy (1862 Barth) ;

4° sous réserves : le **klessem** (*klēsem*), entre Chari et Logone au Sud de Fort-Lamy (1912 Benton d'après Barth) ;

5° le **mousgou** (*musgu*) ou *muzuk* ou *masa*, avec dialecte *luggoy*, sur les deux rives du Logone de Mousgoum en aval jusqu'à Laï en amont (1886 F. Müller d'après Krause) ;

6° le **mandara** ou *wandara* ou *wandala* ou *ndara* ou *alna*, dans la région montagneuse au Sud du Tchad et notamment dans les provinces de Mora et Maroua (1826 Klaproth) ;

7° le **fali** ou *falli* ou *tuburi* ou *tufuri*, dans la région de Toubouri, entre le Logone et la haute Bénoué (1912 Benton d'après Barth) ;

8° sous réserves : le **moundang** (*mundaṅ*) ou *mundā* ou *mbana* ou *lak* ou *laka* ou *laga* ou *yasiṅ*, à l'Est de Garoua sur le haut Mayo-Kebbi (région de Léré et Lamé) et dans quelques enclaves en pays mousgou et en pays sara (1910 Strümpell) ;

9° le **margui** (*margi*), avec dialecte *mulgu* ou *molgoy*, à l'Ouest du mandara dans les régions de Koptchi et Madagari (1912 Benton d'après Barth) ;

10° le **gamergou** (*gamergu*), au Sud de Dikoa dans les provinces

de Douré ou Doré et Dogoumba (1912 Benton d'après Barth);

11° le **karékaré** (*karekare*) ou *kerrikerri*, au Sud-Ouest du Bornou (1905 Merrick);

12° le **bola** ou *bole* (dialectes *fika* ou *pika* et *bara*), parlé par les *Bolawa* (sing. *Bolantši*) à l'Ouest des monts Mandara et au Nord du Baoutchi, dans les régions de Goudjba et Nafada (1854 Koelle);

13° le **guéra** (*gera*), au Sud du précédent (1905 Merrick);

14° le **zani** (dialectes *bullo*, *galla* et *mbutudi*), au Nord de la haute Bénoué, à l'Ouest des Margui (1912 Benton d'après Barth);

15° le **batta**, parlé par les autochtones de l'Adamaoua, notamment à l'Est de Yola dans les régions de Garoua et Mbéré (1910 Strümpell);

16° le **bodé** (*bode*), dialectes *ñgodžin* et *dwéy* ou *doey* ou *deba*, à l'Ouest du Bornou, au Sud du Manga et du Mounio (1849 Clarke et 1854 Koelle);

17° à 30° quatorze langues parlées dans le Baoutchi, région montagneuse comprise entre les vallées du Komadougou au Nord et de la Bénoué au Sud : le **tangalé** (*tañgale*); le **gourka** (*gurka*); le **dougguéra** (*duggera*); le **ouadja** (*wadža*), dialectes *tula* et *awok*; le **bouta** (*buta*); le **rone** (*ron*) (pour ces six langues, 1913 Migeod d'après Hastings, Fitzpatrick, S. Smith et Francis); — l'**ankoué** (*añkwe*), dialectes *montol* et *sura* ou *mažavul*, dans le Baoutchi oriental (1913 Migeod d'après Fitzpatrick, S. Smith et Francis); — le **djaraoua** (*džarawa*), dans la province de Yakoba, Baoutchi central (1854 Koelle); — l'**angas** ou *karañ*, même région (1913 Ormsby); — le **hâm** (*hām*) ou *džabam* ou *fubam* ou *džaba* ou *doma* ou *agato*, même région (1854 Koelle); — le **payem** (1913 Migeod d'après S. Smith); — le **bouroum** (*burum*), Baoutchi occidental (1913 Migeod d'après Francis); — le **koro**, même région, et le **yasgoua** (*yasgũa*), à l'Ouest du Baoutchi et à l'Est du pays gbari (1854 Koelle);

31° le **haoussa** (*hausa*) ou *afno* ou *afnu* ou *kendyi* ou *abakpa* ou *šeše* ou *žãgwe* ou *maraba*, parlé par environ quatre millions de Haoussa habitant les provinces de Tahoua, Zinder, Sokoto, Gober, Zanfara, Kano, Katséna, Gando, Zaria, etc., ou répandus, au milieu d'autochtones de langues diverses, dans le Kebbi, le

Yaouri, le Gbari, le Baoutchi, le Kororofa, l'Adamaoua, etc., ou dans des colonies plus lointaines (oasis du Sud-Algérien, du Sud-Tripolitain et de l'Aïr, centres de Sansanné-Mango, Salaga, Kintampo, Oua et autres dans le Sud-Est de la Bouche du Niger, etc.), ainsi que, comme langue auxiliaire et commerciale, par quantité de Touareg au Sud et au Sud-Est d'Agadès (Kel-Oui notamment) et de nègres habitant les bassins du Komadougou, de la Bénoué, du bas Niger, etc., soit au total par plus de six millions d'individus (1809 Grey Jackson).

BIBLIOGRAPHIE. — Il n'existe pas d'ouvrage d'ensemble sur les langues nigéro-tchadiennes. Pour beaucoup, nous ne possédons que de courts vocabulaires, tels que ceux de la *Polyglotta africana* de KOELLE (London, 1854, in-fol.), ceux que recueillit Barth en 1852 et qui ne furent publiés qu'en 1912 par BENTON (*Notes on some languages of the Western Sudan*, Oxford, in-18), quelques-uns donnés en 1905 par MERRICK dans *Hausa proverbs* (London, in-8), en 1906 par GAUDEFROY-DEMOMBYNES dans *Documents sur les langues de l'Oubangui-Chari* (Paris, in-8), en 1910 par STRÜMPELL dans la *Zeitschrift für Ethnologie* et en 1921 par Sir H. JOHNSTON dans le *Journal of the African Society*, ou de simples numérations rassemblées en 1913 par MIGEOD dans le 2e vol. de *The languages of West-Africa* (London, in-8).

Pour certaines, nous avons des vocabulaires plus étendus et une étude grammaticale dans *Sammlung und Bearbeitung Central-Afrikanischer Vokabularien* (Gotha, 1862-1866, 2 vol. in-8) de BARTH (kotoko, mandara), dans *Die Musuk-Sprache* (Wien, 1886, in-8) de Fr. MÜLLER d'après Krause (mousgou), dans les *Notes on Bolanchi* qui forment la première partie des *Notes on some languages of the Western Sudan* (1912) de BENTON (bola), dans *A manual of the Angas* (London, 1915, in-16) de FFOULKES (angas). Seul le haoussa a fait l'objet de travaux importants et nombreux, dont les meilleurs semblent être *Haussa Grammatik* de J. LIPPERT (Berlin, 1906, in-8) et *Grammaire et contes haoussas* de LANDEROIN et TILHO (Paris, 1909, in-18) et *Dictionnaire haoussa* des mêmes (Paris, 1910, in-18).

X. — *Groupe nigéro-camerounien* (66 langues).

De tous les groupes soudanais et guinéens, le groupe nigéro-camerounien est celui qui renferme le plus grand nombre de langues distinctes. Il est possible d'ailleurs qu'une connaissance plus approfondie des idiomes de la Cross-River et du Cameroun

occidental, sur lesquels nous n'avons que des données très insuffisantes, amène à rattacher à une même langue divers parlers qu'il a paru convenable ici de dissocier, mais le total des langues du groupe n'en demeurera pas moins singulièrement élevé, surtout dans la partie orientale de son domaine.

Celui-ci comprend, d'une façon générale, les basses vallées de la Bénoué et du Niger, la haute vallée de la Sanaga-Djerem, la vallée de la Cross-River ou rivière de Calabar et les vallées des cours d'eau côtiers situés entre le delta du Niger et Porto-Novo. Il est limité au Nord par les groupes nigéro-tchadien et voltaïque, à l'Est par les groupes chari-ouadien et oubanguien, au Sud par le groupe bantou — avec lequel la démarcation est très difficile à établir nettement — et par la mer (à part la petite enclave constituée sur le delta du Niger par le groupe bas-nigérien), à l'Ouest par le groupe éburnéo-dahoméen.

Les populations qui parlent des langues nigéro-camerouniennes appartiennent exclusivement à la race nègre. Elles sont pour la plupart très frustes, mais quelques-unes d'entre elles ont atteint un assez haut degré de civilisation et se sont fait remarquer soit par l'invention d'une écriture (Bamom), soit par une industrie artistique remarquable (Edo), soit par des institutions politiques développées et des aptitudes commerciales dignes d'attention (Yorouba).

Des classes nominales à préfixes existent dans toutes les langues du groupe, à des stades divers. Le système semble se présenter à un état plus complet dans les langues du Cameroun, de la Cross-River et du bas Niger que dans celles de la Bénoué, où cependant les préfixes de classe sont facilement reconnaissables. Dans certaines langues (dioukoun, gayi, dama, mounchi, yakoro, aboua, gbogolo, fi, kana, biobolo, gbira, gara, yala, kpoto, bo, gori, cha, do, sobo, yorouba), les pronoms de classe paraissent avoir disparu, ainsi que la préfixation de l'indice de classe au qualificatif. Dans d'autres (ki, koï, kèlé, sopon, kouri, etc.), chaque classe a conservé son pronom spécial, tant comme pronom proprement dit que comme déterminatif, bien qu'il existe à chaque nombre

un pronom commun pouvant être employé pour des noms de n'importe quelle classe. Dans d'autres langues (ndé, kparabon, gbaragba, kouni, diba, kpé, kounakouna, kayon, sosso, etc.), l'accord en classe du qualificatif avec le nom qualifié, par le moyen d'un préfixe de classe, semble bien attesté.

Le préfixe de classe accolé au nom sert à distinguer celui-ci du verbe. Toutefois il existe dans quelques langues des noms sans préfixe, qui ne se différencient des verbes que par la place respective donnée aux uns et aux autres dans la phrase.

Conjugaison : à l'affirmatif, l'aoriste et l'injonctif sont constitués par le simple radical verbal, les autres aspects verbaux étant généralement marqués chacun tantôt par un préfixe, tantôt par un suffixe spécial ; au négatif, chaque aspect verbal a son préfixe ou son suffixe. Parfois on use du redoublement du radical (futur affirmatif en diba et en kouri). Il arrive souvent que les suffixes se placent, non pas immédiatement après le radical verbal, mais à la fin de la proposition. Enfin, dans un certain nombre de langues (kouri, sopon, mom, gori, ndé, ki, etc.), la négation peut s'exprimer par un simple changement de ton, accompagné ou non d'une modification de la voyelle radicale. Le nombre est quelquefois marqué dans les verbes par une modification de la consonne initiale de la racine, qui passe de l'occlusive nasalisée au singulier à la constrictive correspondante au pluriel, ou s'il n'y a pas de constrictive correspondante, à l'occlusive simple (gbaragba, kparabon).

Ordre des mots : le nom ou pronom sujet précède le verbe, sauf qu'en bo le pronom sujet de la 1re pers. du sing. suit généralement le verbe ; — le nom ou pronom complément d'un verbe suit ce verbe ; — le nom ou pronom complément d'un nom suit ce nom, avec ou sans préfixation au complément tantôt d'une véritable particule d'annexion (comme *ti* en yorouba), tantôt d'un pronom représentant le nom complété (ki : *ba-ne-t ba d'i* « les gens ceux [de] moi, mes gens ») ; le plus souvent, dans les noms composés, le préfixe de classe du nom complément disparaît (ki : *ẹ-tše* « œuf », *ẹ-kwa* « poule », *ẹ-tše-kwa* « œuf de poule ») ; — le qualificatif précède le nom qualifié que l'on ne veut pas

déterminer, mais suit le nom qualifié que l'on tient à déterminer [1] ; c'est dans ce dernier cas surtout que, dans beaucoup de langues du groupe, le qualificatif prend le préfixe de classe du nom qualifié ou un indice d'autre forme mais de même classe (ki : *a-kǫp ba-ne* « de grandes personnes », *ba-ne ba-kǫp* ou *ba-ne-t ba-kǫp* « les gens grands ») ; — le déterminatif (souvent constitué par le pronom de classe quand il y a détermination simple) suit le nom déterminé ; — enfin le nom de nombre suit le nom de la chose nombrée, tantôt en demeurant invariable et alors le nom de la chose nombrée peut rester au singulier, tantôt en s'accordant en classe avec ce dernier nom qui, alors, se met au pluriel (ki : *bo-tõ bo-bǫñe* « une oreille », *a-tõ a-fè* « deux oreilles », *a-tõ a-tañe* « cinq oreilles ») ; quelquefois cependant, le nom de nombre, considéré comme un véritable substantif ayant pour complément le nom de la chose nombrée, précède ce dernier (aboua : *o-tu* « maison », *i-yala l-o-tu* « paire de maisons, deux maisons).

Toutes les langues du groupe semblent posséder des tons musicaux à valeur tantôt étymologique et tantôt grammaticale.

Les désinences des mots sont le plus souvent vocaliques ; cependant, surtout dans les langues de la Cross-River et du Cameroun, on entend fréquemment des désinences consonantiques qui proviennent pour la plupart de l'addition au nom du déterminatif *k*, *t* ou *p*.

Le pluriel des noms se marque exceptionnellement par l'addition d'un suffixe de pluralité (*-ka* ou *-ge* en dioukoun). La règle le plus généralement suivie — et, la plupart du temps, uniquement suivie — consiste à changer le préfixe de classe selon le nombre : « être humain » *u-ndi* plur. *be-ndi* (dama), *u-ne* plur. *be-ne* (gayi), *o-ne* plur. *a-ne* (yakoro), *o-ne* plur. *ba-ne* (ki), *n-ne* plur. *a-ne* (koï et ndé), *nu-ne* plur. *ba-ne* (kparabon), *u-ni* plur. *a-ni* (gbaragba), *o-ni* plur. *a-ni* (kouni), *o-nõ* plur. *a-nõ* (sopon), *o-no* plur. *a-no* (kounakouna), *o-nen* plur. *be-nen* (kayon), *a-nǫ* plur. *ba-nǫ* (kpoto), *o-n'e* plur. *i-n'e* (bo), *ǫ-ni* plur. *e-ni* (yorouba), etc.

1. Parfois cependant l'adjectif précède un nom déterminé, mais il faut alors que celui-ci soit accompagné d'un déterminatif (aboua : *a-nono* « oiseau », *e-bi a-nono* « un bel oiseau », *e-bi a-nono-k* « le bel oiseau »).

Suivent les formes essentielles des pronoms en quelques langues du groupe :

	Dioukoun	Ki	Koï	Ndé	Kparabon
1re pers. sing.	*mi*	*me, d'i*	*me*	*me*	*me*
— plur.	*ba*	*bę, te*	*wot, ra*	*wu*	*so, de*
2e pers. sing.	*u*	*wo, yen*	*wa, ya*	*wo, yę*	*wa, ya*
— plur.	*ne*	*bęn*	*wun*	*na*	*wo, wuna*
3e pers. sing. commun	*a*	*ę*	*we*	*a*	*a*
— plur. —	*nene*	*bę*	*a*	*a*	*ę*

	Gbaragba	Yala	Bo	Sosso	Do	Yorouba
1re pers. sing.	*me*	*mi*	*mu*	*me*	*me*	*mi*
— plur.	*wur*	*ma*	*in'i*	*ma*	*ma*	*wa*
2e pers. sing.	*wo, ya*	*wo*	*i, gi*	*u, ę*	*u, we*	*wo*
— plur.	*wun*	*wa*	*unu*	*wa*	*wa*	*ęyi*
3e pers. sing. commun	*e*	*o*	*ya, o*	*a*	*o*	*ę*
— plur. —	*ba*	*a*	*fa, a*	*a*	*a*	*a*

Les 66 langues dont paraît se composer le groupe peuvent se répartir ainsi :

1° à 7° sept langues parlées sur la rive droite de la Bénoué, dans les provinces du Mouri ou Hamaroua, du Kororofa et de Nassaraoua, et dans la partie du Mouri qui déborde sur la rive gauche : le **dioukoun** (*d'uku*) ou *d'uku* ou *d'ukõ* ou *džukun* ou *kororofa* ou *gbagbã* ou *apa* ou *appa* ou *urapã* ou *kwana* ou *koana* (1849 Clarke et 1854 Koelle) ; — le **boritsou** (*boritsu*) (1854 Koelle) ; — le **yergoum** (*yergum*), appelé aussi *appa* comme le dioukoun (1828 Kilham) ; — le **regba** ou *eregba*, le **mbariké** (*mbarike*) et le **bassa-bénoué** (*basa*, à ne pas confondre avec le « bassa-niger » du même groupe), avec le dialecte *kamuku* (1854 Koelle) ; — le **kagoro**, dialectes *tumu* et *ndob* ou *ndov* ou *burukem* (1849 Clarke et 1854 Koelle) ;

8° à 13° six langues parlées au Sud des précédentes, sur la rive gauche de la Bénoué : le **foudou** (*fudu*) ou *afudu* ou *afuru* et le **gayi** ou *alege* (1854 Koelle) ; — le **dama** (1914 Thomas) ; — le **mounchi** (*munši*) ou *mitši* ou *tiwi*, appelé aussi *appa* comme le dioukoun et le yergoum (1849 Clarke et 1854 Koelle) ;

— le **yakoro** (1914 Thomas) ; — sous réserves le **nkoum** (*ṅkum*) (néant) ;

14° à 24° onze langues parlées au Sud de l'Adamaoua : le **vouté** (*vute*) ou *bute* ou *mbam* ou *mbum*, région de Ngaoundéré, haut Djerem et haute Sanaga (1854 Koelle) ; — sous réserves le **kapoulla** (*kapulla*), région de Tibati (néant) ; — le **lou** (*lu*) ou *balu* ou *bali* ou *pakot*, près et à l'Est du précédent (1849 Clarke et 1854 Koelle) ; — le **tikar**, au Sud-Ouest de Tibati (1880 von Bary) ; — sous réserves le **sagba** ou « sarhba », au Sud du précédent (1912 Gehr) ; — le **koum** (*kum*) ou *kumbe* ou *bakum* ou *bankon* ou *bo* (à ne pas confondre avec le bo ou ibo du bas Niger), entre Tibati et Foumbân (1849 Clarke et 1854 Koelle) ; — le **gba** ou *bagba*, au Nord du précédent (1854 Koelle) ; — le **nio** (*n'o*) ou *ban'o* ou *ban'ō* ou *bayō* ou *bayu̦* ou *bañki̦* ou *pati*, dans la province de Banyo, au Nord-Ouest de Tibati (1849 Clarke et 1854 Koelle) ; — le **mom** ou *bamom* ou *bamum* ou *bamun*, dans la province de Foumbân, au Sud-Ouest du précédent, possédant une écriture propre (1849 Clarke sous le nom de *baipa* et 1854 Koelle) ; — le **pé** (*pe*) ou *pape* ou *papu* ou *papia*, à l'Ouest de Foumbân, et le **gha** (*ga*) ou *gam* ou *bagam* ou *paẓam* ou *bałap* ou *egap*, à l'Ouest du précédent, écrit au moyen des caractères mom (1849 Clarke et 1854 Koelle) ;

25° à 44° vingt langues parlées le long de la Cross River, de la région de ses sources à celle de son embouchure : le **fout** (*fut*) ou *fot* ou *mfut* ou *bafut*, dialectes *penin* ou *binin*, *mbe*, *nẓo* ou *banẓo* et *fum* ou *bafum*, au Nord et au Nord-Ouest de Foumbân (1826 Prichard, 1828 Kilham, 1849 Clarke et 1854 Koelle) ; — le **gouala** (*gwala*) ou *ṅgwala*, au Nord-Ouest de Foumbân (1849 Clarke sous le nom de *barikan* et 1854 Koelle) ; — le **ménia** (*men'a*) ou *momen'a*, près du précédent (1854 Koelle) ; — le **kongouan** (*koṅgwā*) ou *akwoṅgo* ou *banene*, au Nord-Est d'Old-Calabar, le **ki** ou *ṅki* ou *boki* ou *osikom* (dialectes *basua* et *banda* ou *bendega*), au Nord de la haute Cross-River, et le **koï** (*koy*) ou *ekoy* ou *itun* ou *atam* ou *akwa* ou *moko* ou *indẓor* ou *eafen*, au Sud de la haute Cross-River (1849 Clarke et 1854 Koelle) ; — le **kparabon** (*kparàbō*) ou *akparabō*, rive droite, au Sud du ki

(1907 Dayrell) ; — le **ndé** (*nde*) ou *mbeñkpe*, appelé aussi *atam* ou *otam* comme le koï, dialectes *befu* ou *mbofõ*, *afunatam* ou *ekamtulufu* et *akadzu*, rive droite, en aval du précédent (1849 Clarke et 1854 Koelle) ; — le **kèlé** (*kele*) ou *ukele*, rive droite, au niveau du coude nord-ouest de la rivière (1849 Clarke sous le nom de « tshari » et 1914 Thomas) ; — le **nkodo** (*ñkodo*), en arrière du précédent, et le **gbaragba** ou *agbaragba*, rive gauche, en face du kparabon (1914 Thomas) ; — le **kouni** (*kuni*) ou *okuni* ou *udom* ou *obam*, dialectes *ikom* et *olulomo*, entre le gbaragba et le koï (1849 Clarke et 1854 Koelle) ; — le **sopon** (*sopõ*) ou *asopõ* ou *esopõ* ou *okam* ou *wakande*, dialectes *ndadza-nawe*, *arun* ou *adun* et *igbo* ou *imaban* ou *afo*, à l'intérieur du coude nord-ouest, avec quelques enclaves sur la rive droite (1854 Koelle) ; — le **diba** ou *ediba*, rive gauche, en aval du précédent, et le **kouri** (*kuri*) ou *ekuri*, dialectes *ñkpani* et *ge* ou *ugep*, à l'Est du diba et au Sud du sopon (1914 Thomas) ; — le **kounakouna** (*kunakuna*) ou *akunakuna* ou *akurakura*, dialectes *abini* et *umon*, deux rives, en aval du diba (1849 Clarke et 1854 Koelle) ; — le **kpé** (*kpe*) ou *akpet*, à l'Est du précédent (1914 Thomas) ; — le **oué** (*we*) ou *uwet*, au Sud du kouri, et le **kayon** (*kayõ*) ou *akayõ* ou *okon'õ*, rive gauche, en aval du kounakouna (1874 Goldie) ; — le **fi** ou *efik* ou *ibibio* ou « calabar », appelé aussi *moko* comme le koï (dialectes *kwa*, *mona*, *tinan* ou *anañ*, *ikot*, etc.), à cheval sur l'embouchure de la Cross-River, depuis Old-Calabar jusqu'au delta du Niger (1828 Kilham sous le nom de *karaba* et 1846 Waddell) ;

45° à 60° seize langues parlées dans la vallée du bas Niger : le **gbari** ou *gbali* ou *gwali*, avec le dialecte *musu*, provinces de Zaria et de Keffi, au Nord du confluent de la Bénoué et du Niger (1849 Clarke sous le nom de « tshamba » et 1854 Koelle) ; — le **noupé** (*nupe*) ou *nufe* ou *nife* ou *tappa* (dialectes *ebe* ou *agalati*, *kupa*, etc.), dans le Noupé propre (région de Bida) et sur les deux rives du Niger depuis les rapides de Boussa jusqu'au confluent de la Bénoué (1828 Kilham, 1849 Clarke et 1854 Koelle) ; — le **bassa-niger** (*basa*), rive droite du Niger à l'Ouest du noupé (1854 Koelle) ; — le **kakanda** ou *kakandza* ou *sabe* ou *tšabe*

ou *šabi* ou *ad'akai'e*, près du précédent (1837 Laird et Oldfield); — le **gbira** ou *igbira* ou *koto* ou *kotokori* (dialectes *panda*, *bima* et *igu*), sur la rive droite de la Bénoué près de son confluent et sur la rive droite du Niger en aval dudit confluent jusqu'à Idah, et le **gara** ou *igara* ou *igala* ou *okpoto* (distinct du « kpoto » ou « okpoto » mentionné plus loin), à l'intérieur du coude formé par la basse Bénoué et le Niger, rive gauche (1849 Clarke et 1854 Koelle); — le **yala**, dispersé en trois îlots (*agala* entre le gara et le yakoro, *yala* du Nord entre le mounchi et le ndé et *yala* du Sud près et au Nord d'Okouni, haute Cross-River) (1854 Koelle); — le **kpoto** ou *okpoto*, dans le Nord-Est du pays des Ibo (1914 Thomas); — le **bo** ou *ibo* ou *iswama* (dialectes *isele* ou *is'ele*, *abadza* ou *abwadza*, *aro*, *mbof'a* ou *ekpaf'a*, *ala*, *atoma*, *ubuluku*, *nsukwa*, *abi*, *asaba*, *okwuzi*, *boni*, *ndoki*, *ñgwa*, *asa*, *omoma*, *afikpo*, *ezzi*, *mbo*, *ñkanu*, *oweri*, *okugba*, *ñkaraya*, *biko*, *adunsoba*, *ezza*, etc.), sur les deux rives du Niger en aval d'Idah jusqu'au delta exclu et, à l'Est, jusqu'à la vallée de la Cross-River exclue (1827 et 1828 Kilham, 1841 Norris); — le **gbogolo** ou *ogbogolo*, sur la branche orientale du delta du Niger, l'**aboua** (*abwa*), même région, le **kana** ou *ogoni* ou *kereka*, à l'Est du delta, et le **biobolo** ou *ebiobolo* ou *andoni* ou *toridoni*, au Sud du précédent, sur la côte, entre le fi et le djo (1914 Thomas); — le **gori**, à l'Ouest du Niger, entre le kakanda et le sosso (1914 Thomas); — le **sosso** (*soso*) ou *kukuruku* (dialectes *wepa*, *wano*, etc.), rive droite du bas Niger en aval du gbira (1910 Thomas); — le **cha** (*ša*) ou *eša* ou *išā* ou *ibewe*, en aval du précédent (1849 Clarke sous le nom de « nago n°104 » et 1854 Koelle);

61° à 66° six langues parlées le long de la côte du Bénin et en arrière : le **do** ou *edo* ou *bini* ou « bénin » (dialectes *egbele* et *oloma*), province de Bénin (1827 et 1828 Kilham, 1849 Clarke et 1854 Koelle); — le **sobo** (dialectes *igabo* ou *igabor* et *ubobo*), au Sud du précédent, à l'Ouest du bo et au Nord du djo (1828 Kilham et 1848 Allen et Thomson); — le **zékiri** (*zekiri*) ou *izekiri* ou *dzekri* ou *šekri* ou *okiri*, sur la côte entre la rivière Forcados à l'Est et la rivière Bénin à l'Ouest (1828 Kilham et 1837 Laird et Oldfield); — le **yébou** (*yebu*) ou *idzebu*, sur la côte, à

l'Ouest du précédent (1841 Norris et 1845 d'Avezac) ; — le **yorouba** (*yoruba*) ou *yariba* ou *ayo* ou *eyo* ou *egba* ou *nago* ou *ayadzi* ou *aku* (dialectes *ota*, *idzesa*, *ife*, *yagba*, *eki*, *dzumu* ou *akandza* ou *akuya* ou *abinu*, *oworo* ou *egbe* ou *eyagi*, *ondo* ou *doko*, *akoko*, *holli*, etc.), parlé sur la côte depuis la rivière Bénin à l'Est jusqu'à l'Ouémé à l'Ouest (région de Lagos), en arrière jusque vers le 8° lat. nord (régions d'Abéokouta, Ibadan, Ilorin, etc.) et, plus à l'Ouest, au Nord du Dahomey propre (1827 et 1828 Kilham, 1829 Clapperton, 1831 Raban, 1841 Norris, 1843 Crowther) ; — l'**ana** ou *atakpame*, en arrière du mina et de l'éhoué (1902 P. Fr. Müller).

BIBLIOGRAPHIE. — Il n'existe pas d'ouvrage d'ensemble sur les langues nigéro-camerouniennes, mais beaucoup d'entre elles sont représentées dans divers recueils de vocabulaires, tels que la compilation sans valeur de CLARKE (*Specimens of dialects*, London, 1849, in-8), les vocabulaires imparfaits mais encore très utiles de KOELLE (*Polyglotta africana*, London, 1854, in-fol.), les *Dialects of the Okuni district* de DAYRELL (Lagos, 1907, in-16), l'excellent recueil de Northcote THOMAS (*Specimens of languages from Southern Nigeria*, London, 1914, in-4). Quelques-unes ont fait l'objet de travaux spéciaux, parmi lesquels on peut citer : pour le dioukoun, W. K. FRASER, *Vocabulary of the Jukon language*, Zungeru, 1908, in-16 ; pour le même et le mounchi, E. DAYRELL, *Vocabulary of Juku and Munshi*, Zungeru, 1908, in-16 ; pour le tikar, un vocabulaire de VON BARY dans *Zeitschrift der Gesellschaft für Erdkunde zu Berlin*, Band XV, 1880 ; pour le koum, F. SPELLENBERG, *Die Sprache der Bo oder Bankon in Kamerun*, Berlin, 1922, in-8 ; pour le kouni, DAYRELL, *Ikom folk-stories*, London, 1903, in-4 ; pour le fi, GOLDIE, *Principles of the Efik grammar*, 2d ed., Edinburgh, 1868, in-12, et *Dictionary of the Efik language*, Glasgow, 1874, in-8 ; pour le gbari, EDGAR, *A grammar of the Gbari language*, Belfast, 1909, in-8 ; pour le noupé, CROWTHER, *Elements of a grammar and vocabulary of the Nupe language*, London, 1864, in-8: pour le gbira, COOMBER, *Igbira primer and vocabularies*, London, 1869, in-8, et WILLIAMS, *Reading-book in the Igbira language*, London, 1883, in-12 ; pour le gara, COOMBER, *Igara primer and exercises*, London, 1867, in-8, et BYNG-HALL, *English-Okpolo vocabulary*, Zungeru, s. d., in-16 ; pour le bo, SPENCER, *An elementary grammar of the Ibo language*, London, 1901, in-12, et Northcote W. THOMAS, *Anthropological report on the Ibo-speaking peoples*, London, 1913-1914, 6 vol. in-8 ; pour le sosso, du même auteur, le 2e vol. de *Anthropological report on the Edo-speaking peoples*, London, 1910, in-8, et STRUB, *Essai d'une grammaire de la langue kukuruku*, dans *Anthropos*, 1915 ; pour le do, les vol. I et II de l'ouvrage précité de THOMAS sur *the Edo-speaking peoples* ; pour le yorouba, JACQUOT, *Etudes sur la langue nago ou yorouba*, Lyon, 1880,

in-8 : pour l'ana, P. Fr. MÜLLER, *Ein Beitrag zur Kenntniss des Atakpame*, dans *Zeitschrift für Afrikanische, Ozeanische und Ostasiatische Sprachen*, 1902.

XI. — *Groupe bas-nigérien* (1 langue).

Le delta du Niger est occupé par une population nègre, celle des Idjo, qui présente cette particularité de parler une langue qu'il ne semble pas possible de rattacher au groupe au sein duquel elle forme une enclave ni à aucun des autres groupes négro-africains. Il convient donc, au moins provisoirement, de considérer cette langue comme constituant à elle seule un groupe, lequel a quelques points communs avec le groupe nigéro-camerounien qui l'environne, mais s'en sépare nettement par ailleurs, de même qu'il se rapproche, mais à certains égards seulement, du groupe éburnéo-libérien et de certaines langues du groupe voltaïque.

Bien qu'il ne semble pas y avoir de pronoms de classe ni d'accord en classe du qualificatif avec le nom, l'existence de classes nominales paraît certaine. Mais celles-ci sont caractérisées, tantôt par des préfixes vocaliques, tantôt par des suffixes syllabiques, tantôt par les deux procédés à la fois : *ko-mọ* ou *tu-mbwọ* « être humain », plur. *a-ke-mẹ* ou *a-to-ma* ; *o-we-bọ* « homme », plur. *o-wi-ma* ; *i-ọrọ-bọ* « femme », plur. *ẹrẹ-ma*.

Ces affixes nominaux distinguent le nom du verbe ; mais ils peuvent faire défaut et, alors, la place respective du nom et du verbe différencie seule ces deux catégories de mots.

La conjugaison semble procéder tantôt par préfixes et tantôt par suffixes.

Le nom ou pronom sujet précède le verbe ; le nom ou pronom complément d'un verbe *précède* ce verbe, se plaçant entre le sujet et l'expression verbale : *ay o-biri be-famo* « je chien suis-frappant, je frappe un chien » ; *o in ẹre* « il moi voit, il me voit » ; le nom ou pronom complément d'un nom *précède* ce nom : *a-la-bọ* « chef », *o-moni-bọ* « esclave », *a-la-b o-moni-bọ* « l'esclave du chef », *in o-moni-bọ* « mon esclave » ; le qualificatif,

le déterminatif et le nom de nombre *précèdent* le nom qualifié, déterminé ou nombré : *dob o-biri* « gros chien », *kal o-biri* « petit chien » ; *te* « arbre », *me te* « l'arbre » ; *oy a-te* « dix arbres ».

La désinence d'un mot est toujours vocalique, mais la voyelle finale s'élide si le mot suivant est à initiale vocalique.

Il existe des tons musicaux à valeur tantôt étymologique et tantôt grammaticale.

Le pluriel des noms se marque, soit par changement, addition ou suppression de préfixe, soit par changement de suffixe, soit par modification de la voyelle radicale ou du ton, soit par plusieurs de ces procédés à la fois : *te* « arbre » plur. *a-te*, *ware* « maison » plur. *a-ware*, *ko-mọ* « être humain » plur. *a-ke-mẹ*, *o-we-bọ* « homme » plur. *o-wi-ma*, *i-ọrọ-bọ* « femme » plur. *ẹrẹ-ma*, *a-ka* (ton haut) « dent » plur. *a-ka* (ton bas).

Les pronoms personnels offrent cette particularité qu'à chaque nombre la 1re personne est distinguée de la 2e par le ton :

1re pers. sing.	*ay, iy, in*	(ton haut),	plur.	*wa, amene*	(ton haut)
2e pers. —	*ẹy, iy, in*	(ton bas),	—	*wo, omene*	(ton bas)
3e pers. —	*o, a, e, or, ar, er,*		—	*in'e, ne.*	

Le **djo** (*dzo*) ou *idzo* ou *id'o* ou *ud'o* ou « bonny » ou « new-calabar » ou *okrika* ou *akrika* ou *akasa* renferme plusieurs dialectes : *degema* (New-Calabar), *okuloma* ou *obane* (Bonny), *nembe* ou *numbe* (Brass et Noun), *dzo* propre (Forcados), etc.

BIBLIOGRAPHIE. — Les premiers renseignements ont été fournis par Köler (*Einige Notizen über Bonny an der Küste von Guinea, seine Sprache und seine Bewohner, mit einem Glossarium*, Göttingen, 1848, in-8). La *Polyglotta* de Koelle renferme deux vocabulaires (*udzo* et *okuloma*). Il faut encore citer : Taylor, *Ijo primer and vocabulary*, London, 1862, in-8 ; F. E. G. Johnson, *Vocabulary of the Bonny language*, Lagos, 1903, in-18 ; un *Primer in the Ijo language*, London, 1911, in-12 ; les vocabulaires de trois dialectes donnés par Northcote W. Thomas dans ses *Specimens of languages from Southern Nigeria*, London, 1914, in-4, etc.

XII. — *Groupe voltaïque* (53 langues).

Le domaine de ce groupe occupe tout le bassin supérieur des diverses branches de la Volta et s'étend à l'Est jusqu'au Niger et même un peu au delà, où il se soude au groupe nigéro-tchadien, tandis qu'à l'Ouest il atteint le Bani. Limité au Sud par les groupes nigéro-camerounien et éburnéo-dahoméen, il se trouve enclavé, sur la majeure partie de ses frontières, à l'intérieur du groupe nigéro-sénégalais. Les populations parlant des langues voltaïques appartiennent exclusivement à la race nègre ; elles sont nombreuses et la plupart offrent une densité de peuplement remarquable. Certaines, comme les Mossi, ont atteint depuis fort longtemps un degré de civilisation très appréciable ; d'autres, comme les Lobi, sont encore barbares et sauvages ; d'autres encore, comme les Sénoufo, sont simplement frustes.

Toutes les langues voltaïques possèdent des classes nominales qui semblent avoir procédé autrefois par préfixation du pronom ou indice de classe au radical du nom, mais qui, aujourd'hui, procèdent en général par suffixation. Certaines cependant, comme le gourma, le kabré, le konko, le bassari, le losso, etc., ont conservé des préfixes de classe, tout en faisant en même temps usage de suffixes, usant souvent à la fois des deux systèmes : « terre » *teñ-ga* (mô), *kẹ-teñ* (konko), *ga-teñ-ga* (gourma). Les pronoms de classe ont subsisté dans certaines langues dites « gourounsi » (nourouma, sissala, etc.), à un degré moindre dans certaines autres (dagari, gbanian), et, comme déterminatifs et relatifs, dans d'autres (tem, kabré, etc.), mais, le plus souvent, il n'y a plus qu'un seul pronom de la 3e personne à chaque nombre. Les adjectifs tantôt s'accordent en classe avec le nom qu'ils qualifient (gourma, konko, kabré, tem, etc.), tantôt demeurent invariables quelle que soit la classe du nom (mô, sénoufo, etc.).

L'affixe de classe distingue le nom du verbe, mais il peut faire défaut et il tombe souvent en composition : alors c'est la place respective du nom et du verbe qui les fait reconnaître.

La conjugaison repose sur un jeu de préfixes ou de suffixes

affirmatifs et de préfixes négatifs, dont chacun caractérise un aspect verbal. On fait également usage d'auxiliaires suffixés pour certains aspects secondaires.

Ordre des mots : le nom ou pronom sujet précède le verbe ; — le nom ou pronom complément d'un verbe suit ce verbe, sauf en sénoufo où, comme dans la plupart des langues nigéro-sénégalaises, il précède le verbe, se plaçant entre le préfixe de conjugaison, s'il existe, et le radical verbal (mô *kõ ko-m* « donne eau », sénoufo *to-mõ ka* « eau donne », « donne de l'eau ») ; — le nom ou pronom complément d'un nom précède ce nom ; dans les noms composés, le suffixe de classe du nom complément disparaît et, s'il s'agit de langues faisant usage de préfixes de classe, le nom composé prend le préfixe du nom complété (mô *teñ-ga* « terre », *sõ-ba* « maître », *teñ-ga sõ-ba* « le détenteur d'un terrain », *teñ-sõ-ba* « un chef de terre » ; gourma *ga-teñ-ga* « terre », *o-dan-o* « chef », *o-teñ-dan-o* « un chef de terre ») ; — le qualificatif, le déterminatif et le nom de nombre suivent le nom qualifié, déterminé ou nombré.

Les désinences des mots sont uniquement vocaliques ou nasales.

Il ne paraît pas exister de tons musicaux dans la plupart des langues du groupe ou du moins ils semblent n'y avoir que peu d'importance ; dans d'autres, et notamment dans celles du sous-groupe de l'Est, ils se rencontrent au contraire avec une valeur très nette, tantôt étymologique et tantôt grammaticale.

Les gutturales labialisées *gb* et *kp* sont fréquentes, ainsi que les sons *g* et ɣ, qui sont en général interchangeables ; il existe en sénoufo une laryngale rappelant le « ʿaïn » arabe. Dans certaines langues, comme le tem, la voyelle du pronom sujet et parfois celle du préfixe de conjugaison varient selon la voyelle radicale du verbe, de même que la voyelle du suffixe nominal peut varier selon la voyelle radicale du nom et inversement ; dans les mêmes langues, la sourde initiale de la racine verbale ou nominale devient sonore après le pronom sujet du verbe ou complément du nom non suivi d'un déterminatif (tem *ma wa* pour *ma fa* « que je donne », *e di-nde* pour *e ti-nde* « son poisson »).

Le pluriel des noms se marque par un suffixe de classe appro-

prié. Dans les langues faisant usage de préfixes, le préfixe disparaît au pluriel ou se modifie comme le suffixe (sénoufo *ñ'e-ge* « œil » plur. *ñ'e-le*; mbouin *bi-o* « enfant » plur. *bi-ma* ; dagari *bi-le* « enfant » plur. *bi-we* ; mampourou *bi-a* « enfant » plur. *bi-si* ; mô *bī-ga* « enfant » plur. *bī-si* et *mō-ga* « un Mô » plur. *mō-si* ; tem *tem-ne* « un Tem » plur. *tem-ba* ; bassari *kẹ-bi-kẹ* « enfant » plur. *m-bi-am* et *ko-di* « maison » plur. *te-di-te* ; gourma *ga-bi-ga* « enfant » plur. *gi-bi-ge* et *li-d'ẹ-li* « maison » plur. *d'ẹ-na*).

Voici les formes essentielles des pronoms dans quelques langues :

	Gourma	Tem	Mô	Mbouin	Sénoufo
1re pers. sing.	*mi*	*ma*	*m*	*mi*	*mi*
— plur.	*ti*	*da*	*d*	*si*	*su*
2e pers. sing.	*fi*	*n'a*	*f*	*bi*	*ma, mu*
— plur.	*i*	*mi*	*i*	*mō*	*mu, ye*
3e pers. sing.	*we*	*a, wa*	*a*	*wo*	*u*
— plur.	*ba*	*ba*	*ba*	*ba*	*pe*

A noter que plusieurs langues voltaïques semblent posséder à la 1re personne du pluriel un pronom inclusif et un pronom exclusif.

Au point de vue linguistique, le domaine des parlers voltaïques est loin d'être complètement exploré. Dans l'état actuel de nos connaissances, le groupe paraît renfermer 53 langues, qu'il conviendrait de répartir en sept sous-groupes :

1° à 11° sous-groupe de l'Est : le **kambari** ou *kambali* ou *yauri*, deux rives du Niger en amont de Boussa (1854 Koelle) ; — le **gourma** (*gurma*) ou *bi*, dans le royaume des Gourmantché ou Bimba ou de Fada-n-Gourma, 150.000 individus environ (1849 Clarke et 1854 Koelle) ; — le **ber** ou *berba* ou *barba* ou *bariba* ou *bergo* ou *bargu*, dans la région du Bergo ou Borgou (1854 Koelle) ; — le **kabré** (*kabre*) ou *kabere* ou *kaure* ou *kauri* (dialectes *sum* ou *sumba*, *basila*, *pilapila*, *dompago* et *wēdẓiwēdẓi*), dans la région des monts Atakora et au Sud (1828 Kilham, 1849 Clarke et 1854

Koelle) ; — le **legba** ou *debba* dans les provinces de Djébiga et Pama et à l'Est de Sansanné-Mango (1849 Clarke et 1854 Koelle) ; — le **konko** (*kon̄ko*) ou *kon̄kobiri* ou *kon̄komba* ou *t'opo* ou *t'opowa*, à l'Ouest des monts Atakora (1911 Groh) ; — le **tem** ou *tim* ou *temba* ou *k'amba* ou *brinni* ou *kotokoli* ou *tšautšo*, entre Sansanné-Mango et Sokodé (1849 Clarke et 1854 Koelle); — le **kassélé** (*kasele*) ou *akasele* ou *tsamba* ou *tšamba* ou *t'āsi*, région de Tchamba, entre le haut Mono et Sokodé (1768 Oldendorp) ; — le **mouâ** (*mwā*) ou *mōba*, Nord-Ouest du Togo, (1911 Groh); — le **losso** (*loso*), dans le district de Sokodé, et le **bassari** (*basari*) ou *tobote*, dans le district de Bassari (1911 Groh) ;

12° à 23° sous-groupe de la Volta Blanche : le **mô** (*mō*) ou *mōrẹ* ou *mōle* (nom de la langue) ou *mōẓo* ou *mōgo* (nom du pays) ou *mōsi* ou *mōse* (nom des habitants), parlé par les Mossi des royaumes du Yatenga et de Ouagadougou et par un certain nombre d'étrangers, 1.600.000 individus environ au total (1819 Bowdich, 1849 Clarke et 1854 Koelle) ; — le **nankana**, à l'Est et au Sud-Est de Léo jusque vers Gambaga (1912 Tauxier) ; — le **djélagna** (*dẓelan̄a*) ou *tšelan̄a*, région de Sansanné-Mango (1854 Koelle) ; — le **boura** (*bura*), au Sud du nankana (1912 Tauxier) ; — le **mampourou** (*mampuru*) ou *mampursi* ou *mamprusi*, région de Gambaga (1911 Groh, Migeod) ; — le **koussan** (*kusā*) ou *kusāsi*, au Sud-Est du précédent (1913 Westermann) ; — le **kandia** (*kand'a*) ou *kand'aga*, au Nord-Est de Oua (1849 Clarke sous le nom de « tshamba n° 207 » et 1854 Koelle sous le nom de *guresa*) ; — le **dagomba** ou *dagboma* ou *dagbane*, régions de Tamalé et Yendi, à l'Est de la Volta Blanche (1819 Bowdich et 1849 Clarke) ; — le **gbanian** (*gban'ā*) ou *gban'e* ou *gwān* ou *gwāndẓa* ou *gondẓa* ou *gwan'a* ou *bāndẓa* ou *ban'a* (appelé aussi *nta* ou *bōre* ou *bōle* comme une langue éburnéo-dahoméenne de la même région), dans les provinces de Salaga, Daboya et Bôlé (1819 Bowdich sous le nom d'*ingwa*) ; — sous réserves, le **oule** (*ule*) ou *wule* ou *wulewule*, sur la Volta Noire, au Nord du 11° lat. nord, avec quelques colonies plus au Sud (néant) ; — le **dagari** ou *dagarti*, deux rives de la Volta Noire au Nord et au

Sud de Oua, où se parle le dialecte *wala* (1904 Delafosse) ; — le **birifo** ou *birifõ* ou *birifor*, à l'Ouest de la Volta Noire du 10° au 9°5' lat. nord environ (1904 Delafosse) ;

24° à 31° sous-groupe dit « gourounsi » : le **kourouma** (*kuruma*) ou *fulse* ou *n'on'ose*, 80.000 individus environ, autochtones du Yatenga (1917 Tauxier) ; — sous réserves le **kô** (*kõ*) ou *kipirsi*, au Sud du précédent (néant) ; — le **nourouma** (*nuruma*) ou *nunuma* ou *nibulu* ou *guresi* ou *grusi* ou *gurunsi* ou *grunši*, avec le dialecte *menk'era*, au Sud du précédent, à l'Est de la Volta Noire et à hauteur de Boromo (1849 Clarke) ; — le **kasséna** (*kãsena*) ou *kasene* ou *kasuna* ou *kasum* ou *kasm* ou *kãsem* ou *kãsom* ou *kãsom-bura*, à l'Est de Léo (1854 Koelle) ; — le **fra** ou *frafra* ou *kãsom-fra* ou *k'ãlo* ou *ak'ulo* ou *adžolo* ou *yulu* ou *le* ou *re*, ville de Léo et région au Sud de cette ville (1854 Koelle) ; — le **sissala** (*sisala*) ou *isala* ou *kwama* ou *bagbalã*, au Sud de Léo (1854 Koelle); — le **siti**, village de Vonkoro, sur la Volta Noire à hauteur de Bouna (1904 Delafosse) ; — le **dégha** (*dega*) ou *mõ* (ne pas confondre avec le « mô » ou « mossi ») ou *g'amu* ou *d'ammu* ou *buru*, trois villages, dont Assafoumo, au Nord-Est de Bondoukou (1904 Delafosse) ;

32° à 35° sous-groupe dit « lobi » : le **pougouli** (*puguli*) ou *buguri*, sur le Bougouriba à l'Ouest de Diébougou ; le **dian** (*d'ã*) ou *žã* ou *žãga*, canton de Diébougou ; le **lobi**, 100.000 individus environ, région de Gaoua et plus à l'Ouest, ainsi qu'au Nord-Ouest de Bouna ; le **gan** (*gã*), district de Lorhosso ou Lokhosso (pour ces quatre langues, 1904 Delafosse) ;

36° à 43° sous-groupe dit « koulango » : le **koulan** (*kulã*) ou *ṅgorã* ou *kulãṅgo* (pays) ou *kulãmvo* (habitants) ou *kpaẓala* ou *pakalla*, régions de Bouna et de Bondoukou (1819 Bowdich sous le nom de *gaman* et 1849 Clarke sous le nom de *boutuku*) ; — le **lorhon** (*loẓõ*) ou *loẓo* ou *loẓoma* ou *logoma* ou *nabe* ou *nẹmbay*, à l'Ouest du précédent le long de la Comoé (1921 Tauxier) ; — le **tégué** (*tege*) ou *teges'e* ou *tuna* ou *tumbe*, au Sud de Lorhosso (1921 Tauxier d'après Labouret) ; — sous réserves : le **padorho** (*padoẓo*), à l'Ouest de Lorhosso ; le **dorhossié** (*doẓos'e*), haute Comoé ; le **komono**, au Nord-Est de Kong ; le **karaboro**, région

de Lorhognilé; le **kiéfo** (*k'efo*) ou *t'efo*, au Sud de Bobo-Dioulasso (pour ces cinq langues, néant) ;

44° à 52° sous-groupe dit « bobo » : le **boua** (*bwa*) ou *bobofi* (dialectes *sãkura* et *n'ẹnege*), 100.000 individus environ, dans la boucle de la Volta Noire (1912 Tauxier) ; — le **tara** ou *pwe* ou *boboule*, 110.000 individus environ, au Nord-Ouest du précédent (1849 Clarke et 1904 Delafosse) ; — le **kian** (*k'ã*) ou *bobogbẹ*, 40.000 individus environ, à l'Ouest de la haute Volta Noire (1904 Delafosse) ; — le **mbouin** (*mbwẽ*) ou *kpẽ* ou *turuka* ou *turka*, 50.000 individus environ, régions de Léra, Banfora et Bérégadougou (1909 Tauxier, inédit) ; — sous réserves : le **tagba**, à l'ouest de Bobo-Dioulasso ; le **nanergué** (*nanerge*), le **vigué** (*vige*), le **toussia** (*tus'a*) et le **sémou** (*semu*), cercle de Bobo-Dioulasso (pour ces cinq langues, néant) ;

53° sous-groupe ne comprenant que le **sénoufo** (*senufo*) ou *s'ena* ou *s'ene* (dialectes *bamãna* ou *min'ãñka*, *seneẹẹ* ou *sendere*, *nobolo* ou *nabulu*, *zõna*, *kadle*, *põmporo*, *n'ene*, *teneure*, *foro* ou *folo*, *gbãto*, *kasembele*, *kãfibele*, *kofolo*, *k'embaẹa*, *nafaẹa*, *gbãnzoro*, *n'aẹa-folo*, *pala* ou *kpalaẹa* ou *pallaka*, *sikolo*, *tafile* ou *tafire*, *falafala*, *takpọnẽ* ou *tagbọna* ou *tagwana*, *gimini*, *d'ammala*, *nafãna* ou *pãn-tara* ou *bãnda*, etc.), parlé dans les régions de Koutiala, Sikasso, Odienné, Boundiali, Niellé, Korhogo ou Koroko, Darhakolon-dougou, Dabakala, Kong, Bondoukou, etc., du Bani au coude sud de la Volta Noire, par 800.000 individus environ (1887 Tautain et 1904 Delafosse).

BIBLIOGRAPHIE. — Les langues du groupe voltaïque ont fait l'objet d'esquisses d'ensemble de la part de DELAFOSSE (*Mémoires de la Société de Linguistique de Paris*, 1911) et de WESTERMANN (*Anthropos*, 1913, et *Zeitschrift für Kolonialsprachen*, 1913-1915). De plus, deux chapitres (VI et VII) leur sont consacrés dans : DELAFOSSE, *Vocabulaires comparatifs de plus de 60 langues ou dialectes*, Paris, 1904, in-8. Sur quelques-unes d'entre elles ont été publiés des travaux spéciaux, par exemple : sur le gourma, le bassari, le kassélé et le ber, D. WESTERMANN, pp. 36 à 142 de *Die Sprache der Guang... und fünf andere Togosprachen*, Berlin, 1922, in-8; sur le tem, P. Fr. MÜLLER, *Beitrag zur Kenntnis der Tem-Sprache* (*Mitteilungen des Seminars für Or. Spr.*, 1905) ; sur le mô, F. FROGER, *Étude sur la langue des Mossi*, Paris, 1910, in-8, et *Manuel pratique de langue môré*, Paris, 1923, in-8 ; sur le dagomba, R. FISCH,

Grammatik der Dagomba-Sprache, Berlin, 1912, in-8 ; sur le kasséna, J. CREMER, *Grammaire kasséna*, Paris, en cours de publication ; sur le dian, J. HAILLOT, *Étude sur la langue dian* (*Bulletin du Comité d'études historiques et scientifiques de l'A.O.F.*, 1920) ; sur le sénoufo, G. CHÉRON, *Essai sur la langue minianka* (*ibid.*, 1921).

XIII. — *Groupe éburnéo-dahoméen* (48 langues).

Ce groupe fait suite à l'Ouest, le long de la côte du golfe de Guinée, au groupe nigéro-camerounien ; il est limité au Nord par l'extrémité occidentale de ce dernier et par le groupe voltaïque, au Nord-Ouest par le groupe nigéro-sénégalais et à l'Ouest par le groupe éburnéo-libérien, qu'il rencontre sur le bas Bandama ; il possède en outre une enclave isolée à l'extrémité occidentale de ce dernier groupe, dans le Libéria, enclave constituée par le gola. Parmi les langues qu'il renferme, plusieurs sont extrêmement voisines les unes des autres, comme le fanti et le tchi, l'agni et le baoulé. Les langues de ce groupe sont parlées, les unes par des populations assez arriérées, la plupart par des peuples remarquablement doués au point de vue intellectuel ou artistique, appartenant tous d'ailleurs à la race nègre.

Toutes les langues éburnéo-dahoméennes ont possédé autrefois un système de classes nominales à préfixes, qui fonctionne encore aujourd'hui à l'état parfait ou presque parfait dans quelques-unes, localisées dans des domaines extrêmement restreints, telles que le logba, l'adélé, le kabou, le kposso, le kédémonié, le trougbou, le balé, le béri, l'adioukrou, tandis que, le plus souvent, il n'en demeure que des vestiges sous forme de préfixes nominaux plus ou moins généralisés.

Le nom est facilement distingué du verbe dans les langues où le procédé des classes nominales a subsisté dans son intégrité. Ailleurs il arrive que les préfixes nominaux eux-mêmes font défaut, et alors c'est la place respective du nom et du verbe qui, seule, sert à les différencier.

Les divers aspects verbaux sont marqués à l'affirmatif par des

préfixes ou, plus rarement, par des suffixes, l'aoriste étant toujours constitué par le radical seul et l'injonctif par le radical simple ou nasalisé à l'initiale. La négation s'indique à l'injonctif par un préfixe spécial, aux autres aspects par un préfixe, ou, plus souvent, par une particule négative suffixée soit au verbe, soit à la proposition, parfois par un changement de ton (kédémonié, trougbou, balé, goua) ou par la nasalisation de la consonne initiale du radical verbal (akan, fanti, tchi, abron, zéma, agni, vétéré, abouré, goua) ou par le redoublement de cette consonne (attié). Plusieurs langues possèdent un soi-disant passif caractérisé par un préfixe *a-* qui paraît être de nature pronominale (comparez le soi-disant passif du haoussa et d'autres langues nigéro-tchadiennes).

Ordre des mots : le nom ou pronom sujet précède le verbe ; — le nom ou pronom complément d'un verbe suit ce verbe, parfois en perdant son préfixe (avikam, aladian, abè) ; en gola toutefois, le *pronom* complément du verbe se place entre le préfixe de conjugaison (lequel est alors en réalité un verbe auxiliaire) et le radical verbal ; — le nom ou pronom complément d'un nom précède ce nom, soit directement, soit en intercalant entre les deux une particule d'annexion ; le nom complété peut perdre son préfixe ; dans les langues à classes nominales intactes, le nom complément perd souvent aussi son préfixe pour former un nom composé, lequel prend alors en général le préfixe de classe du nom complété (kédémonié : *ǫ-kukǫ* « poule », *li-dze* « œuf », *ǫ-kukǫ li-dze* « l'œuf d'une poule », *li-kukǫ-dze* « un œuf de poule ») ; parfois, comme en fon, en balé, etc., le pronom complément du nom est remplacé par un véritable adjectif possessif qui, comme tout adjectif, suit le nom (fon : *xwe to-we* « maison tienne, ta maison ») ; — le qualificatif, le déterminatif et le nom de nombre suivent le nom qualifié, déterminé ou nombré.

La désinence d'un mot est toujours vocalique ou nasale, sauf en adioukrou. Dans les autres langues et parfois même dans celle-ci, lorsque la racine se termine par une consonne, on en forme un mot en lui suffixant une voyelle généralement identique à la voyelle radicale.

Des tons musicaux à valeur soit étymologique soit grammaticale existent dans quelques langues du groupe (fon, éhoué, kédémonié, trougbou, balé, goua, etc.), mais, dans la plupart des autres, leur existence semble problématique.

Le pluriel des noms se marque par une modification du préfixe de classe (kédémonié *o-nu* « être humain » plur. *ba-nu*, *ku-pi* « poil » plur. *si-pi* ; béri *ku-du* « dizaine » plur. *a-du* ; tchi *o-hene* « chef » plur. *a-hene* ; vétéré *e-bra* « femme » plur. *m-bra* ; goua *o-tõ* « bœuf » plur. *n-tõ* ; adioukrou *lẽ-ten* « pirogue » plur. *mẽ-ten*) ou bien il est indiqué par l'addition d'un suffixe de pluralité (fon *xwe* « maison » plur. *xwe-le*, baoulé *swä* « maison » plur. *swa-mõ*, abouré *k'a* « homme » plur. *k'a-me*). Parfois on a au pluriel addition d'un préfixe, ou bien d'un préfixe et d'un suffixe, ou modification de la désinence (tioko *bara* « femme » plur. *m-bara-m*, *buru* « dizaine » plur. *a-bru* ou *a-bro* ; baoulé *ba* « enfant » plur. *ma* (pour *m-ba*) ou *ma-mõ*, *buru* « dizaine » plur. *a-bura*).

Voici les pronoms personnels dans quelques-unes des langues du groupe :

	Fon	Kédémonié	Trougbou	Balé	Tchi
1re pers. sing.	*ñ'ẹ, m*	*mẹ*	*mẹ*	*mi*	*me, mi*
— plur.	*mi* (ton haut)	*blọ*	*blọ*	*bu*	*ame*
— plur. excl.	.....	.....	.....	.....	*ye*
2e pers. sing.	*a, we*	*wo*	*wọ*	*fo*	*wo*
— plur.	*mi* (ton bas)	*mbọ*	*wọnọ*	*bi*	*amene*
3e pers. sing.	*e*	pr. de classe	*yẹ*	*a*	*o*
— plur.	*ye*	pr. de classe	*ba*	*ba*	*be*

	Baoulé	Attié	Abouré	Adioukrou	Gola
1re pers. sing.	*me, mi*	*mẹ*	*me, mi*	*me*	*me, mi*
— plur.	*ame*	*a*	*ame*	*se*	?
— plur. excl.	*ye*	.....	*e*	*we*	*wi*
2e pers. sing.	*wo, e*	*bọ*	*wo*	*e, i*	*mbo*
— plur.	*amụ*	*mun*	*ama*	?	?
3e pers. sing.	*o, a*	*o, e*	*o, e*	*o*	pr. de classe
— plur.	*be*	*ba*	*ve*	*be*	pr. de classe

Les 48 langues du groupe peuvent se répartir en sept sous-groupes :

1° à 5° sous-groupe dit « éhoué » : le **mahi** (*maxi*), au Nord d'Abomey (1849 Clarke et 1854 Koelle) ; — le **fon** (*fõ*) ou *fõgbe* ou *dãxọme* ou *dẓẽdẓi*, Dahomey propre et régions d'Allada et Ouidah (1730 Père Labat) ; — le **mina** ou *gẽ* ou *gẽgbe* ou *popo* ou *watši* ou *anexo*, régions de Grand-Popo et Anecho, sur la côte (1827 et 1828 Kilham) ; — le **krépé** (*krepe*) ou *añfwe* ou *kpãndo*, régions de Misahöhe et Kpando (1819 Bowdich, 1849 Clarke et 1854 Koelle) ; — l'**éhoué** (*ehwe*) ou *ewe* ou *efe* ou *eve* ou *yewe* ou *wegbe* ou *añlo* ou *awuna*, régions d'Avrékété, Lomé et Quittah, sur la côte (1856 et 1857 Schlegel) ;

6° à 24° sous-groupe de la Volta : le **logba**, au nord et au Nord-Est d'Avatimé (1903 Westermann) ; — l'**adélé** (*adele*), région d'Adélé, à l'Est du bas Oti (1895 Christaller) ; — le **kabou** (*kabu*) ou *kẹbu* ou *akabu* ou *kögbörikö*, au voisinage du précédent (1907 Wolf) ; — le **kposso** (*kposo*) ou *akposo*, dans la région des monts Akposso (1907 P. Fr. Müller) ; — le **kédémonié** (*kedemon'e*) ou *avatime*, région d'Avatimé (1887 Christaller) ; — le **balé** (*bale*) ou *santrokofi*, parlé par 1500 individus seulement entre la Volta et les monts Akposso (1911 Funke) ; — le **trougbou** (*trugbu*) ou *n'ãgbo-tafi*, parlé dans treize villages entre Avatimé et Kpando (1910 Funke) ; — le **bouem** (*bwem*) ou *lefana*, entre le kposso au Nord et le krépé au Sud (1910 Westermann) ; — l'**ago** ou *ogo* ou *bogo* ou *ahlõ* ou *axolo*, près des sources du Dayi, dans le moyen Togo (1898 Plehn-Seidel) ; — sous réserves : le **likpé** (*likpe*), à l'Ouest du précédent ; l'**akpafou** (*akpafu*), dans le moyen Togo ; le **bôli** (*bõli*) ou *bowli*, même région ; le **boro**, parlé à Tapa et à Ouoraouora, même région (pour ces quatre langues, néant) ; — l'**adan** (*adã*) ou *adãpe* ou *adãgbe* ou *adãme*, avec le dialecte *krobo*, région d'Addah, à l'embouchure de la Volta (1854 Koelle) ; — le **gan** (*gã*) ou *ga* ou *akra*, ville et région d'Accra, sur la côte (1788 Isert) ; — le **gouan** (*gwã*) ou *gwañ* ou *akripõ*, cantons d'Anoum, de Laté, de Cherepong, etc., coude de la basse Volta (1844 Prichard) ; — le **fétou** (*fetu*) ou *afutu* ou *awutu* ou *obutu* ou *gomwa* ou *dwoma* ou *aguna*, région de Winnebah, sur la côte, à l'Ouest d'Accra (1480 Eustache de la Fosse, publié en 1897, et 1675 W. J. Müller) ; — l'**oti** ou *nawuri*, vallée

du bas Oti (1922 Westermann) ; — le **nta** ou *inta* ou *gwan'a* ou *bōle* ou *daboya* (ne pas confondre avec le gbanian, du groupe voltaïque, dit aussi « nta, gouania, bôlé » et parlé dans la même région), provinces de Bôlé, de Daboya, du Gondja, et plus au Sud (1819 Bowdich) [1] ;

25° à 31° sous-groupe dit « tchi », englobé avec les deux suivants sous le nom de *tō* par les Mandingues : le **béri** (*beri*), village de Taghadi, à l'Ouest de la Volta Noire, entre Bouna et Bondoukou (1921 Tauxier); — le **koranza** ou *ñkoranza*, région de Kintampo (1849 Clarke sous le nom de « dagamba n° 289 ») ; — le **kouahou** (*kwahu*) ou *okwahu* ou *borō* ou *brō* ou « abron de l'Est », à l'Est et au Sud-Est du précédent dans les provinces d'Ataboubou et d'Amina (1819 Bowdich sous le nom de *burum* et 1849 Clarke sous les noms de *trubi* et « quako ») ; — l'**akan** (*akā*), dialectes *akwambu*, *akwapim* et *akim*, au Sud du précédent (1828 Rask) ; — le **fanti** (*fānti*), dialectes *asin*, *t'efo*, *wasa* ou *wasaw* et *fānti* propre ou *agwa* ou *amina*, provinces de Cape-Coast, Elmina, Sekondi, et le pays en arrière (1764 Protten) ; — le **tchi** (*tši*) ou *twi* ou *t'i* ou *k'i* ou *otši* ou *ok'i* ou *okin* ou *asānti*, parlé par les Adansi, les Denkira, les Amansi, les Assanti ou Achanti et les Ahafo, région de Coumassie (1819-1821 Hutton) ; — l'**abron** (*abrō*) ou *brō* ou *borō* ou *abonō* ou *abonu* ou « abron de l'Ouest » ou *g'amā* ou *gamā* ou *boga* (dialectes *abrō* propre, *doma* ou *domna*, *ntakima*, etc.), région à l'Est, au Sud-Est et au Sud de Bondoukou (1849 Clarke sous le nom de « boutuku n° 171 ») ;

32° et 33° sous-groupe dit « apollonien » : l'**ahanta**, provinces d'Axim et Dixcove près du cap des Trois-Pointes (1819 Bowdich) ; — le **zéma** (*zema*) ou *nsima* ou *zimba* ou *amanaya*, province de Béyin, basse Tano, et colonies plus à l'ouest, notamment à Grand-Bassam (1819 Bowdich) ;

34° à 37° sous-groupe dit « agni » : le **tioko** (*t'oko*) ou *t'okosi*

1. Il existerait encore, au Togo, d'autres langues, en voie de disparition, qu'il y aurait lieu sans doute de ranger avec les précédentes, mais sur lesquelles nous n'avons aucune documentation, telles que celles des Tétémang, des Lolobi, des Bayika.

ou *tšokosi*, isolé dans une colonie d'origine baoulé à Sansanné-Mango et aux environs (1911 Groh) ; — l'**assayé** (*asaye*) ou *sahïve* ou *sęfïvi*, à l'Ouest de la moyenne Tano (1904 Delafosse) ; — l'**agni** (*añi*) ou *añ'i* ou *aowim* ou *awõñïvi* ou *bonnai* (dialectes *bini*, *kumwenu*, *bonna* ou *bonda*, *sikãsu*, *nden'e*, *bet'e*, *bures'a* ou *brusa*, *aris'ẽ* ou *ariš'ẽ*, *afema*, *sãmvį* ou *sãwi*, *asini* ou *asoko*, etc.), au Sud de l'abron de l'Ouest, entre l'assayé et le fanti à l'Est et la Comoé à l'Ouest (1714 Père Loyer) ; — le **baoulé** (*baule*) ou *poni* (dialectes *ñg'e* et *abe* du Mango, *bomo*, *ndamę*, *wure* ou *wor'e* ou *wele*, *moro* ou *moronu*, *agbęñ'aų*, *baule* propre, *g'asale* ou *t'asale*, etc.), entre la Comoé et le Bandama (1898 Lasnet) ;

38° à 47° sous-groupe dit « des lagunes » : l'**attié** (*at'e*) ou *ak'ę* ou *kurobu* (dialectes *nedẽ* et *bodẽ* ou *budę*), entre la basse Comoé et l'Agnéby (1900 Dreyfus) ; — le **vétéré** (*vetere*) ou *v'etre* ou *b'etri* ou *ewutre* ou *ewutile* ou *papaire* ou *mek'ibo*, îles et rives des lagunes Tano, Abi, etc. (1904 Delafosse) ; — l'**abouré** (*abure*) ou *abonwã* ou *akaples*, provinces de Bonoua à l'Est de la basse Comoé et d'Abra et Mouossou près de son embouchure (1902 Bailleul) ; — le **goua** (*gwa*) ou *ñgora* ou *mbãto* ou *potu*, lagune Potou, et l'**ébrié** (*ebrie*) ou *k'ama*, rive nord de la lagune Ebrié entre la lagune Potou et l'Agnéby et village de Petit-Bassam (1904 Delafosse) ; — l'**abè** (*abę*) ou « abbey », entre l'Agnéby et le bas Bandama et le bas Nzi ; l'**ari** ou *abigi* ou *abidži*, région de Bessédi, au Sud du précédent ; l'**adioukrou** (*ad'ukru*) ou *ag'ukru* ou *og'ukru*, au Nord de la lagune Ebrié à l'Ouest de l'Agnéby, région de Dabou-Toupa ; l'**aladian** (*alad'ã*) ou *alag'ã* ou « jack-jack », au Sud de la lagune Ebrié à l'Ouest de Petit-Bassam, région de Jacqueville (pour ces quatre langues, 1901 Delafosse) ; — l'**avikam** ou *avikǫm* ou *avekwǫm* ou *gbãnda* ou *briñã* ou *kwakwa* ou *labu*, sur la côte, de Krafi à l'Est à Dibou à l'Ouest et sur le bas Bandama, région de Lahou (1849 Leighton Wilson) ;

48° le **gola** ou *gora* ou *gura*, isolé sur le bas Saint-Paul en amont de Monrovia (1849 Clarke et 1854 Koelle).

BIBLIOGRAPHIE. — Aucun travail d'ensemble n'a été publié sur le groupe

éburnéo-dahoméen dans son entier, mais on peut citer, pour les sous-groupes de l'Est : J. G. CHRISTALLER, *Die Volta-Sprachen Gruppe* (*Zeitschrift für Afrikanische Sprachen*, 1887) ; le même, *Sprachproben aus dem Sudan* (*ibid.*, 1889) ; B. GROH, *Sprachproben aus zwölf Sprachen des Togohinterlandes* (*Mitteil. des Seminars für Or. Spr.*, 1911) ; — et pour les sous-groupes tchi, apollonien, agni et des lagunes, les chapitres III et I de DELAFOSSE, *Vocabulaires comparatifs de plus de 60 langues ou dialectes*, Paris, 1904, in-8.

Parmi les travaux consacrés à des langues spéciales du groupe, on peut citer : pour le fon, DELAFOSSE, *Manuel dahoméen*, Paris, 1894, in-18, et Père JOULORD, *Manuel français-dahoméen*, Lyon, 1907, pet. in-8 ; pour le mina, Père SCHUH, *Vocabulaire gengbe-français et français-gengbe*, Rome, 1910, in-12 ; pour l'éhoué, D. WESTERMANN, *Grammatik der Ewe-Sprache*, Berlin, 1907, in-8 ; pour le logba, même auteur, *Die Logbasprache* (*Zeit. für Afr. Spr.*, VII, 1903) ; pour l'adélé, J. G. CHRISTALLER, *Die Adelesprache* (*ibid.*, I, 1895) ; pour le kabou, F. WOLF, *Grammatik der Kögböriko-Sprache* (*Anthropos*, 1907) ; pour le kposso, même auteur, *Grammatik der Kposo-Sprache* (*ibid.*, 1909) ; pour le kédémonié, E. FUNKE, *Versuch einer Grammatik der Avatimesprache* (*Mitteil. des Seminars für O. Spr.*, 1909) ; pour le balé, même auteur, *Die Santrokofisprache* (*ibid.*, 1911) ; pour le trougbou, même auteur, *Die Nyangbo-Tafi-Sprache* (*ibid.*, 1910) ; pour le bouem, D. WESTERMANN, *Die Lefánasprache in Togo*, Berlin, 1910, in-8 ; pour l'ago, même auteur, *Die Ahlö-Sprache in Mitteltogo*, pp. 5 à 33 de *Die Sprache der Guang*, Berlin, 1922, in-8 ; pour le gan, J. ZIMMERMANN, *A grammatical sketch of the Akra or Gã language*, Stuttgart, 1858, 2 vol. in-8 ; pour le gouan, D. WESTERMANN, *Die Sprache der Guang*, Berlin, 1922, in-8 ; pour l'akan, H. N. RIIS, *Elemente des Akwapim-Dialektes der Odschi-Sprache*, Basel, 1853, in-8 ; pour le tchi et le fanti, J. G. CHRISTALLER, *A grammar of the Asante and Fante language*, 2d ed., Basel, 1882, in-8 ; pour l'agni et le baoulé, DELAFOSSE, *Essai de manuel de la langue agni*, Paris-1901, in-8 ; pour l'attié, Père MÉRAUD, *Essai sur la langue attié*, Dabou, 1902, in-18 ; pour le gola, D. WESTERMANN, *Die Gola-Sprache* (*Zeit. für Eingeborenensprachen*, 1921).

XIV. — *Groupe nigéro-sénégalais* (36 langues).

Par son étendue territoriale, ce groupe vient en troisième ligne après le groupe bantou et le groupe nilo-tchadien. Il fait suite à l'Ouest au groupe nigéro-tchadien, contourne au Nord, à l'Ouest et au Sud le groupe voltaïque, poussant des antennes entre ce dernier et les groupes éburnéo-dahoméen et nigéro-camerounien, puis limite au Nord et à l'Ouest le groupe éburnéo-libérien et s'entremêle ensuite au groupe sénégalo-guinéen, atteignant la

mer près de Monrovia, de Freetown, de Conakry, de Bathurst, pour se rencontrer, à la lisière méridionale du Sahara, avec le domaine de l'arabe et du berbère. Il convient d'ajouter que cette dernière langue forme des enclaves importantes, avec divers dialectes parlés par les Touareg, entre Agadès et Gao, à l'intérieur de la Boucle du Niger et près de Tombouctou, de même que l'arabe entre Tombouctou et le Sénégal ; de plus le peul, du groupe sénégalo-guinéen, est parlé en divers points du territoire nigéro-sénégalais, notamment dans le Liptako (Boucle du Niger) et dans le Massina (au Sud-Ouest de Tombouctou).

Les populations qui parlent des langues nigéro-sénégalaises appartiennent toutes à la race noire ; mais les plus septentrionales (Songoï, Azer, Sarakollé) sont plus ou moins métissées, depuis une époque paraissant ancienne, d'éléments de race blanche, sémitiques ou libyco-berbères. A part les tribus du Sud qui vivent au voisinage de la forêt dense et dont quelques-unes sont anthropophages (Dan, Manon), à part aussi quelques tribus isolées comme les Dogon, les Samo, etc., la plupart des peuples de langues nigéro-sénégalaises ont atteint depuis plusieurs siècles un état de civilisation relativement avancé, et certains (Sarakollé, Mandingues, Songoï) possèdent un passé historique remarquable.

Les classes nominales ont complètement disparu, semble-t-il, des langues de ce groupe, quoiqu'on en retrouve des vestiges sous forme d'anciens préfixes de classe, communs à nombre de langues négro-africaines, qui ont subsisté çà et là en s'incorporant, pour ainsi dire, au radical de certains noms, tels le préfixe *mu-* de la classe humaine dans le mot mandingue *mu-so* « femme » $\sqrt{so}$.

En général, le nom ne se distingue du verbe que par la place respective de l'un et de l'autre dans la phrase. En réalité, il n'y a pas de verbes à proprement parler, mais seulement des noms qui, soit par le moyen d'affixes de conjugaison, soit simplement de par la place qu'ils occupent, font fonction de verbes. Cependant, il arrive parfois qu'on distingue le rôle nominal d'un mot de son rôle verbal en lui donnant, quand il est nom, une désinence

spéciale (*o* dans certains dialectes mandingues et dogon, *e* en sarakollé, en azer, en kouranko) ou en nasalisant sa consonne initiale (dialectes bambara et dioula du mandingue).

La conjugaison présente un nombre assez considérable d'aspects verbaux, tant affirmatifs que négatifs, marqués chacun par un préfixe spécial ou, exceptionnellement, par un suffixe, tandis que le parfait affirmatif des verbes intransitifs comporte au contraire le plus souvent un suffixe. En dehors de ce dernier cas, l'absence de complément direct est la seule chose qui distingue le verbe intransitif du verbe transitif et elle suffit à donner à un verbe quelconque une valeur neutre ou passive (mandingue : *à bi muso bugo* « il frappe une femme », *a bi bugo* « il est frappé »).

Le sujet, nom ou pronom, précède le verbe. — Le nom ou pronom complément d'un verbe ou d'un nom précède ce verbe ou ce nom ; il n'y a d'exception qu'en songoï, où le complément du verbe suit en général le verbe ; le zerma, extrêmement voisin du songoï par ailleurs, suit la règle générale, laquelle consiste à mettre le complément du verbe immédiatement avant le radical verbal ; parfois cependant, en zerma, le complément d'un verbe non pourvu de préfixe suit ce verbe au lieu de le précéder. — Le qualificatif, le déterminatif et le nom de nombre suivent toujours le nom qualifié, déterminé ou nombré ; l'expression mandingue *o muso* « cette femme », *o kẹ* « cet homme », ne fait pas exception à la règle, car *o* est un pronom complément du nom qui le suit, et la traduction exacte serait « femme de ceci, homme de ceci ».

Il n'existe pas de désinences proprement consonantiques, sauf en songoï. Dans les autres langues, et souvent même en songoï, lorsque la racine se termine par une consonne, on en forme un mot en lui suffixant une voyelle qui, en général, est identique ou analogue à la voyelle radicale. Commune dans le Sud, la gutturale labialisée *gb* passe généralement à *g* dans le Nord. On rencontre r uvulaire (ɼ), interchangeable avec *g*, dans plusieurs langues ou dialectes, non dans d'autres ; il en est de même de la jota (*x*), qui est interchangeable avec *k*.

On n'a signalé de tons musicaux, avec quelque apparence de

certitude, que dans une seule langue du groupe, le loko, où ils peuvent être dus au voisinage de langues sénégalo-guinéennes à tons.

Les noms forment leur pluriel par l'addition d'un suffixe à la forme du singulier.

Suivent les pronoms personnels dans quelques langues du groupe :

	Zerma	Songoï	Dogon	Sia	Sor-ko	Sara-kollé	Man-dingue	Soussou
1re pers. sing.	*ay*	*ay*	*mi, i*	*mi*	*ni*	*n'e*	*ne, ni*	*ni*
— plur.	*iri*	*yer*	*emę*	?	?	*o*	*an, nu*	*woñ, muxu*
2e pers. sing.	*ni*	*ni*	*u*	*bi*	*ã*	*an*	*i, e*	*i*
— plur.	*wor*	*wor*	*ey*	?	?	*axa*	*al, au*	*wo*
3e pers. sing.	*a, ñga*	*a, ñga*	*a, wo*	*a*	*a*	*a*	*a*	*a*
— plur.	*i, ñgay*	*i, ñgi*	*be*	*ki*	?	*i*	*u*	*e*

Les 36 langues nigéro-sénégalaises peuvent se répartir en six sous-groupes :

1° et 2° sous-groupe du Nord-Est : le **zerma** ou *dẓerma* ou *ẓaberma*, vallée du Niger en aval d'Ansongo jusqu'à Ilo, ainsi qu'à l'Est dans le Djermaganda et au Sud-Ouest dans le Dendi, 100.000 individus environ (1897 Hacquard et Dupuis) ; — le **songoï** (*soñọy*) ou *soñay* ou *sõgọy* ou *sõgay* ou *kuria* (« kissour » dans Caillé), vallée du Niger en amont de Gao inclus jusqu'à Dienné inclus, y compris Tombouctou, 300.000 individus environ, langue commerciale et politique du Nord de la Boucle du Niger, avec des dialectes aberrants à Agadès (*emgedeẓi*), à Tabalbalet (*balbali*), etc. (XIe siècle El Bekri et 1818 Lyon) ;

3° à 9° sous-groupe de l'Est : le **dogon** (*dogõ*) ou *dogom* ou *dogo* ou *toẓo* ou *tombo* ou *habe* (sing. *kado*), parlé sur les montagnes dites « falaises » des cercles de Bandiagara et de Hombori, ainsi que par les *Deforo* ou *Humbebe*, au pied de ces montagnes, 100.000 individus environ, nombreux dialectes (1849 Clarke et 1921 Arnaud) ; — le **samo** ou *samoẓo* ou *sọmno*, dans le Sud du Yatenga, près du coude nord de la Volta Noire et à l'Est de Sikasso (1819 Bowdich et 1899 anonyme sous le nom de

« mossi ») ; — le **sia** (*s'a*), parlé par les autochtones de Bobo-Dioulasso (1904 Delafosse) ; — sous réserves : le **sembla**, le **ouara** (*wara*), le **natioro** (*nat'oro*) et le **blé** (*ble*), dans le cercle de Bobo-Dioulasso (pour la première de ces quatre langues, néant; pour les trois autres, 1921 Ferréol [inédit]);

10° et 11° sous-groupe du Sud-Est : le **boussa** (*būsa*) ou *bisa* ou *bisã* ou *bisõ* ou *bọko*, entre le ber ou bariba à l'Ouest, le kambari au Nord et le noupé au Sud (1854 Koelle) ; — le **boussan** (*busã*) ou *motsã*, parlé par les Boussansi ou Boussangsé ou Motsantsé au Nord de Gambaga et de Sansanné-Mango (1889 Christaller sous le nom de mozanze[1]) ;

12° à 14° sous-groupe du Nord : le **sorko** ou *sorogo* ou *koroñgoy* ou *boso* ou *bozo*, parlé par les pêcheurs et riverains du Niger et de ses lacs et canaux, de Tombouctou à Dienné (1904 Delafosse) ; — le **sarakollé** (*saraxọlle*) ou *sẹrẹkule* ou *sẹrawuli* ou *soniñke* ou *aswanik* ou *marka* ou *gad'aga* ou *wakore*, parlé sur les deux rives du Sénégal dans la région du Galam ou de Bakel, au Nord dans le Guidimaka et au Nord-Est dans la région de Nioro (XIe siècle El Bekri, XIVe siècle Ibn Batouta, et XVIIe siècle vocabulaire anonyme publié en 1845 par d'Avezac) ; — l'**azer** ou *adẓer* ou *girgãke* ou *masīn*, parlé par les nomades Guirganké et par les sédentaires des oasis de Tichit, Ouadân, Oualata, etc., peut-être aussi par les chasseurs Némadi (1855 Barth) ;

15° à 21° sous-groupe du centre : le **ligbi** ou *ligwi* ou *nigbi* ou *nigwi*, entre le coude sud de la Volta Noire et Bondoukou ; le **noumou** (*numu*), parlé surtout par des artisans dans la région de Bondoukou ; le **huéla** (*hw̃ela*) ou *w̃ela* ou *vw̃ela*, même région (pour ces trois langues, 1904 Delafosse) ; — le **mandingue** ou « mandingo » ou *mandẹ* ou *mandị* ou *manẹ* ou *manị* ou *mali* ou *melli*ou *wãgara* ou *kãga* ou *sõgo* ou *sọẓo*, parlé d'une façon continue depuis le Bani à l'Est jusqu'à la basse Gambie à l'Ouest et depuis le parallèle de Nioro au Nord jusqu'aux sources du Niger et aux approches de la forêt dense au Sud, et, d'une façon dissé-

1. Le boussa et le boussan se rapprochent très étroitement des langues formant le sous-groupe du Sud et semblent en constituer, avec d'appréciables solutions de continuité, le prolongement vers l'Est jusqu'au bas Niger, en suivant à peu près la zone de culture des colatiers.

minée, dans de nombreuses et populeuses colonies citadines sur le coude sud de la Volta Noire, à Bondoukou et dans la région, en pays bobo (Boromo, Bobo-Dioulasso, etc.), en pays sénoufo (Sikasso, Bong, Koroko, Kong, etc.), ensemble 2.500.000 individus environ dont le mandingue est la langue maternelle, plus 2.000.000 environ qui le parlent comme langue auxiliaire ; idiome de très grande extension, qui tend à se répandre dans toute l'Afrique occidentale ; nombreux dialectes pouvant se ramener à six sections : *d'ula* ou *g'üla* ou *d'ura*, dans des colonies à l'Est du Bani ; *bãmbara* ou *bãmana*, entre Bani et Niger et à l'Ouest du Niger en aval de Bamako, ainsi que dans le Sahel ; « malinké de l'Est » ou mandingue propre, le long du haut Niger en amont de Bamako et entre Niger et Bafing ; « malinké du Nord » ou *xasoñke*, dans la région de Kayes ; « malinké de l'Ouest », sur la Gambie ; « malinké du Sud », aux abords de la forêt dense ; de plus, il existe un dialecte commun dit *kãgbẹ* (langue blanche), au profit duquel se fait l'expansion de la langue (XIe siècle El Bekri, XIVe siècle Ibn Batouta et Ibn Khaldoun, XVIIe siècle vocabulaire anonyme publié en 1845 par d'Avezac) ; — le **kouranko** (*kurãko*), région des sources du Niger (1916 Thomas) ; — le **kono**, au Sud-Ouest du précédent (1854 Koelle) ; — le **vaï** (*vay*) ou *vẹy* ou *karo* ou *karu*, sur la côte, à l'Ouest du Saint-Paul, possède une écriture syllabique spéciale (1828 Kilham et 1849 Koelle) ;

22° à 36° sous-groupe du Sud : le **gbin** (*gbẽ*) ou *gbeyị* ou *mbẽ* ou *guro*, entre Bondoukou et la Comoé, et le **ngan** (*ñgã*), dans le Mango et le Diammala (1904 Delafosse) ; — le **noua** (*nwa*), sur le haut Bandama Rouge (néant) ; — le **mouin** (*mwẽ*) ou *mona* ou *moni*, à l'Ouest du précédent, entre le Oué et le Béré (1849 Clarke sous les noms de « lorangga » et « kangga n° 207 » et 1901 Delafosse) ; — le **lo** ou *guro*, entre Bandama et Sassandra au Sud de Séguéla (1899 Eysséric) ; — le **toura** (*tura*), montagnes du haut Sassandra (néant) ; — le **dan** (*dã*) ou *mẹbe* ou *yabuba* ou langue des « Dioula anthropophages », haut Cavally (1854 Koelle sous le nom de *gĩo*) ; — le **manon** (*manõ*) ou *gõ*, à l'Ouest du précédent (1854 Koelle sous le nom de *mano*) ; — le

guerzé (*gerze*) ou *gberese* ou *gbẹse* ou *kpẹse* ou *kpẹle* ou *kpwesi* ou *pesa* ou « pessy » ou « barline », entre la haute Nuon et le haut Saint-Paul (1827-1828 Kilham, 1849 Clarke et 1854 Koelle); — le **toma** ou *loma* ou *buzi*, au Sud-Ouest de Beyla (1849 Clarke sous le nom de *baru* et 1854 Koelle); — le **gbandi** (*gbãndi*), au Sud du kono et au Nord du vaï (1854 Koelle); — le **mendé** (*mende*) ou *mendi* ou *koso* ou *wuro* ou *komboya*, à l'Ouest du vaï, du gbandi et du kono (1827 Kilham); — le **loko** ou *landoɔo*, entre la Roquelle et la Grande Scarcie, au Nord-Est de Freetown (1854 Koelle); — le **langan** (*lãgã*) ou *sako* ou *d'alloñke* ou *d'aluñka*, langue des autochtones du Foûta-Diallon (1799 Mungo-Park); — le **soussou** (*susu*) ou *soso*, au Sud-Ouest du Foûta-Diallon (1802 Brunton).

BIBLIOGRAPHIE. — Comme travaux d'ensemble sur une partie au moins des langues nigéro-sénégalaises, on peut citer : H. STEINTHAL, *Die Mande-Neger-Sprachen*, Berlin, 1867, in-8, et les chapitres IV et V de DELAFOSSE, *Vocabulaires comparatifs*, Paris, 1904, in-8. Parmi les publications relatives à des langues du groupe : pour le zerma et le songoï, HACQUARD et DUPUIS, *Manuel de la langue soñgay*, Paris, 1897, in-18, et, pour le songoï, DUPUIS-YAKOUBA, *Essai de méthode pratique pour l'étude de la langue songoï*, Paris, 1917, in-8 ; pour le mandingue, DELAFOSSE, *Essai de manuel pratique de la langue mandé*, Paris, 1901, in-8 (dialecte dioula) ; Père SAUVANT, *Manuel bambara*, Alger, 1913, in-18 (dialecte bambara) ; Père ABIVEN, *Dictionnaire français-malinké et malinké-français précédé d'un abrégé de grammaire malinkée*, Conakry, 1906, in-18 (dialecte malinké de l'Est) ; Ch. MONTEIL, *Les Khassonké*, Paris, 1915, in-8 (dialecte malinké du Nord) ; MACBRAIR, *A grammar of the Mandingo language*, London, 1837, in-8 (dialecte malinké de l'Ouest) ; pour le vaï, S. W. KOELLE, *Outlines of a grammar of the Vei language*, new edition, London, 1902, in-8 ; pour le mendé, F. W. H. MIGEOD, *The Mende language*, London, 1908, in-8 ; pour le soussou, J. B. RAIMBAULT, *Dictionnaire français-soso et soso-français*, Rio-Pongo, 1885, in-18.

XV. — *Groupe éburnéo-libérien* (24 langues).

D'étendue restreinte, ce groupe présente une grande unité entre les parlers qui le composent. Son territoire est limité exclusivement à la zone de forêt dense comprise entre le Bandama et le Saint-Paul, avec une pointe s'avançant au Sud-Est, de l'autre côté du Bandama, jusque sur la lagune Ebrié. Les popu-

lations qui habitent ce domaine, frustes et souvent sauvages, parfois anthropophages, appartiennent toutes à la race noire.

Des vestiges très nets de classes nominales à préfixes subsistent sous la forme de quelques préfixes nominaux et surtout sous celle de deux pronoms de la 3[e] personne du singulier, dont l'un est réservé aux êtres humains, le second servant à représenter tous les autres êtres ou objets.

Le nom, en dehors des cas où le préfixe de classe a subsisté, ne se distingue essentiellement du verbe que par la place respective de l'un et de l'autre.

Il existe deux conjugaisons, l'une pour les verbes simples (radicaux ou dérivés), l'autre pour les verbes accompagnés d'une particule séparable. En général l'aoriste, ainsi que, souvent, l'injonctif, ne prend pas d'affixe de conjugaison, tandis que le parfait se marque à l'aide d'un suffixe et les aspects secondaires à l'aide de préfixes, la négation s'indiquant au moyen d'une particule préfixée au verbe ou, quelquefois, suffixée. La forme du pronom sujet varie fréquemment selon les aspects du verbe et sert alors à les déterminer ou à les préciser.

Ordre des mots : le nom ou pronom sujet précède le verbe; parfois il se place entre le préfixe de conjugaison et le radical verbal ; — le nom ou pronom complément d'un verbe suit ce verbe en général dans les propositions principales, se plaçant entre le radical verbal et le suffixe de conjugaison, mais il le précède dans les propositions subordonnées ou négatives, se plaçant entre le préfixe de conjugaison et le radical verbal, ou même quelquefois avant le préfixe de conjugaison ; lorsqu'il y a interrogation, le complément du verbe peut même se placer avant le sujet ; de plus la place du complément varie parfois selon qu'il est nom ou pronom ou, s'il est pronom, selon la personne, ou enfin selon le type de conjugaison (kra : *ā na nų* « nous boire vin de palme, buvons du vin de palme », *gi ā nų na* « viens nous vin de palme boire, viens que nous buvions du vin de palme », *ā se nų na* « nous non vin de palme boire, nous ne buvons pas de vin de palme ») ; — le nom ou pronom complément d'un

nom précède ce nom, directement si c'est un pronom, généralement en intercalant entre les deux une particule pronominale si c'est un nom (kra : *nă kru* « toi village, ton village », *kru ẹ buro* « village lui chemin, le chemin du village », *ble ẹ döbo* « bœuf lui tête, la tête du bœuf ») ; dans les noms composés, les deux éléments se suivent directement, le complément prenant la forme du pluriel (kra : *ble* plur. *bli* « bœuf », *bli-döbo* « une tête de bœuf », *bli-yu* « un petit de bœuf, un veau ») ; — le qualificatif, le déterminatif et le nom de nombre suivent le nom qualifié, déterminé ou nombré.

Il n'y a que des désinences vocaliques.

Les voyelles nasales sont assez rares, mais la nasalisation des consonnes et notamment des aspirées est fréquente. Les labiales gutturalisées abondent.

Il existe des tons musicaux à valeur étymologique et à valeur grammaticale. Le ton haut est affecté, entre autres choses, à la 1re personne et le ton bas à la 2e, la différence de ton suffisant, même en l'absence de tout pronom, à distinguer l'une de l'autre ces deux personnes.

Le pluriel des noms est marqué par une modification de la voyelle finale, ainsi parfois que par un déplacement de l'accent, une modification de la voyelle radicale ou un changement de ton ; de plus, le préfixe de classe, s'il existe, peut disparaître au pluriel (abri : *bre* « bœuf » plur. *bri*, *tu* « arbre » plur. *ti*, *kọü* « maison » plur. *kẹi*, *kubu* « européen » plur. *kubo*, *ń'ō* « être humain » plur. *ń'u* ; né : *m-ble* « bœuf » plur. *bli*, *kókwẹ* « poule » plur. *kokọ́*, *búro* « maison » plur. *bulé*).

Voici les pronoms personnels dans quelques langues du groupe :

	Dida	Né	Abri	
1re pers. sing.	*ĕ, nă, mŏ*	*ĕ, nă, mŏ*	*ĕ, nă, mō*	(1)
— sing.	*ē, nā, mō*	*ē, nā, mō*	*ē, ī, nā, mō*	(2)
3e pers. s. classe humaine	*ọ*	*ọ*	*ọ*	
— s. autre classe	*ā*	*ẹ̄*	*ẹ̄*	
1re pers. plur.	*ā*	*aa*	*aa*	(1)
2e — —	*ā*	*aa*	*aa*	(2)
3e — —	*ā*	*aa*	*aa*	(3)

	Té	Pla	Kra	
1re pers. sing.	*ĕ, nă, mŏ*	*ĕ, nĕ, nă, mŏ*	*ĕ, nă, mŏ*	(1)
2e pers. sing.	*ē, nā, mō*	*ē, nē, nā, mō*	*ē, nā, mō*	(2)
3e pers. s. classe humaine	*o*	*o*	*ọ*	
3e — s. autre classe	*ă, ẹ̆*	*ă*	*ẹ̆*	
1re pers. plur.	*ă*	*ă*	*ă*	(1)
2e — —	*ā*	*ā*	*ā*	(2)
3e — —	*ā, ō*	*wā, ō*	*ā*	(3)

(1) Ton haut. — (2) Ton bas. — (3) Ton moyen.

Les 24 langues du groupe peuvent se répartir en deux sous-groupes :

1° à 8° sous-groupe oriental : l'**ahizi**, sur les deux rives de l'extrémité occidentale de la lagune Ebrié (néant) ; — le **dida** ou *d'ida*, au nord de l'avikam, du Bandama au Rio Fresco (1849 Clarke sous le nom de *wawi*) ; — le **zégbé** (*zegbe*) ou *kwaya*, vallée du Rio Fresco et Yobéhiri (1849 Clarke sous les noms d'*eple* et de « friesco ») ; — le **go** ou *god'e*, région de Kotrou, entre Fresco et Trépoint, et en arrière (1849 Clarke) ; — le **oua** (*wa*) ou *waya* ou *wad'e* ou *bobwa*, au Nord de Daloa (1904 Delafosse) ; — le **bêté** (*bẹte*), de Daloa inclus au Nord jusqu'à Kouati exclu au Sud (1904 Delafosse) ; — le **koua** (*kwa*) ou *kwad'a* ou *kwadre*, sur les deux rives du bas Sassandra de Kouati inclus à Griguiblé exclu (1905 Thomann) ; — le **né** (*ne*) ou *newole* ou *neyo* ou *nihiri*, villages et cantons de Trépoint (*Grabwa*), de Sassandra (*Bokre*) et Drewin (*Kẹbẹ*) (1849 Clarke) ;

9° à 24° sous-groupe occidental : le **ouobê** (*wobẹ*), région de Sémien à l'Ouest du Sassandra et au Sud-Ouest de Séguéla (néant) ; — le **ba** ou *bañwa* ou *zag'e* ou *ñgere* ou *gere*, entre Sassandra et Cavally, au sud du ouobê et du dan (1905 Thomann) ; — le **bakoué** (*bakwe*), au Sud du précédent (1849 Clarke sous le nom de *pori*) ; — le **houané** (*hwane*) ou *hwẽne*, région de Victory, sur la côte, entre Drewin et San-Pedro (1904 Delafosse) ; — le **pia** (*p'a*) ou *p'ẹ* ou *omelokwe*, vallée du San-Pedro (néant) ; — l'**abri** ou *abriwi* ou *abriñ'ō* ou *berebi* ou *aulo*, région des Béréby entre Tahou et Ouappou (1849 Clarke) ; — le **té** (*te*) ou *tewi* ou *tiwa* ou *tepo* ou *horo*, à l'Est du bas Cavally et sur

ses deux rives, au Sud du bakoué et au Nord de l'abri et du pla, régions de Sapo, Paloubo, Graouro, Krépo, Ouampo, Olodio, Ségré ou Sigli, Grabo, Taté, etc. (1849 Clarke) ; — le **pla** ou *plawi* ou *bla* ou *plapo*, district de Tabou (1849 Clarke) ; — le **bâ** (*bā*) ou *bāpo* ou *bābo*, bouche du Cavally et village de Half-Cavalla (1827-1828 Kilham sous le nom d'*appa*) ; — le **padé** (*pade*) ou *padebo*, entre Cavally et Nuon, au Nord du suivant (1906 Johnston) ; — le **gré** (*gre*) ou *grebo* ou *gedebo* ou *krebo* ou *krepo*, sur la côte, du Cavally à Nifou, région de Cape-Palmas (1835 Leighton Wilson et Mrs. Wilson) ; — le **gbé** (*gbe*) ou *sikoṅ*, au Nord de Sinoe (1854 Koelle) ; — le **kra** ou *krawo* ou *krao* ou *krawi* ou *kru* ou *naña* ou *ñana*, sur la côte dite de Krou, de Nifou à Bafou, régions de Nannakrou, Krouba, Settrakrou, Sinoe, etc. (1827-1828 Kilham) ; — le **bassa** (*basa*) ou *gbasa*, de Bafou à Marshall, région de Grand-Bassa (1827-1828 Kilham) ; — le **gui** (*gi*) ou *givi* ou *gibi* ou *kwia* ou « queah », entre la rivière Duqueah et Monrovia (néant) ; — le **dé** (*de*) ou *do* ou *dewoy*, sur la lagune de Monrovia et le bas Saint-Paul (1854 Koelle).

BIBLIOGRAPHIE. — A citer comme études d'ensemble : Fr. Müller, *Die Sprachen Basa, Grebo und Kru*, Wien, 1877, in-8, et le chapitre II de Delafosse, *Vocabulaires comparatifs*, Paris, 1904, in-8. Comme études de détail : pour le né, G. Thomann, *Essai de manuel de la langue néouolé*, Paris, 1905, in-8 ; pour le gré, J. Payne, *Grebo grammar*, New York, 1864, in-12 ; pour le kra, Usera y Alarcon, *Ensayo grammatical del idioma de la raza africana de Ñana por otro nombre Cruman*, Madrid, 1845, in-8 ; pour le bassa, Crocker, *Grammatical observations on the Basa language*, Edina (Liberia), 1844, in-16.

XVI. — *Groupe sénégalo-guinéen* (24 langues).

Ce groupe se trouve resserré entre le groupe nigéro-sénégalais et la mer. Les langues qui le composent, visiblement acculées à la côte par la poussée des grandes migrations venant de l'Est, n'ont chacune qu'un domaine exigu, à l'exception du peul, qui d'ailleurs, par un mouvement en sens inverse, s'est transporté non seulement vers le Sud (Foûta-Diallon), mais aussi vers l'Est et

bien au delà de son territoire primitif, lequel devait se trouver vraisemblablement entre Nioro et Tombouctou avant les immigrations dans le Foûta-Toro.

Les populations parlant des langues sénégalo-guinéennes appartiennent toutes à la race nègre, sauf pour ce qui est des Peuls ou Foulbé propres, lesquels remontent vraisemblablement à une souche sémitique ancienne, mais sont beaucoup moins nombreux que les métis ou noirs purs parlant le même idiome qu'eux. Les populations noires ou métisses de langue peule (Toucouleurs du Foûta-Toro et Foula du Foûta-Diallon) et les Peuls propres, ainsi que les Ouolofs et les Sérères, ont atteint un certain degré de civilisation; les autres peuples parlant des langues sénégalo-guinéennes sont en général au contraire très arriérés.

Toutes les langues de ce groupe possèdent un système de classes nominales, tantôt fonctionnant à l'état parfait (peul, diola, timné, etc.), tantôt montrant des tendances à la simplification (ouolof, sérère, limba, boulom, krim, kissi, etc.). On procède en général par préfixation de l'indice de classe, parfois par préfixation et suffixation (boulom, où le préfixe de classe devient suffixe quand le nom est accompagné d'un déterminatif ; kissi, où les deux procédés coexistent), uniquement par suffixation dans l'une des langues (peul). Souvent, en ouolof par exemple, les pronoms de classe ne sont plus ou presque plus en usage en tant que pronoms proprement dits et ont disparu de la plupart des noms en tant que préfixes nominaux, mais ils ont subsisté comme déterminatifs, comme relatifs et comme préfixes de l'adjectif pour faire accorder celui-ci en classe avec le nom qu'il qualifie (ouolof : *nit*, pour *ku-nit*, « une personne », *nit ku*, pour *ku-nit ku*, « la personne » ou « la personne qui » ; *nit ku-mag* « une personne âgée », *nit ku-mag ku* « la personne âgée » ; comparez en peul, où l'on procède par suffixation : *ne*dd-*o* « une personne », *ne*dd-*o o* « la personne », *ne*dd-*o mo* « la personne qui », *ne*dd-*o maw*-do « une personne âgée », *ne*dd-*o maw*-do *o* « la personne âgée »). Dans plusieurs langues du groupe (peul, ouolof, sérère), les classes sont caractérisées par la nature de la consonne initiale

de la racine, laquelle, pour une même racine, est, selon les classes, tantôt une constrictive, tantôt l'occlusive simple et tantôt l'occlusive nasalisée correspondante.

L'indice de classe distingue aisément le nom du verbe. Là où il est tombé, la distinction est encore facile pour les noms déterminés ; mais les noms indéterminés ne se différencient des verbes que par la place respective qu'occupent les uns et les autres.

Il existe en général au moins deux voix (active et réfléchie) et souvent trois (active, réfléchie et passive), dont chacune est caractérisée par des suffixes de conjugaison spéciaux, tant au négatif qu'à l'affirmatif. Il n'est jamais fait usage de préfixes, au moins pour les aspects principaux du verbe. En peul, sauf dans le dialecte du Foûta-Diallon et quelques autres, et en sérère, le pluriel des verbes est caractérisé par la substitution, à l'initiale de la racine, de l'occlusive nasalisée à la constrictive.

Le nom ou pronom sujet se place avant le verbe ; pourtant il y a quelquefois inversion du pronom sujet dans les propositions incidentes ou relatives (peul, ouolof, sérère). Le nom ou pronom complément d'un nom suit ce nom. Le qualificatif suit le nom qualifié. Le déterminatif, y compris le pronom de classe employé comme déterminatif, suit le nom déterminé ou, s'il y a aussi un qualificatif, ce qualificatif ; mais s'il est employé avec une valeur nettement démonstrative, il précède le nom (peul : *put'-u ñgu* « le cheval », *ñgu put'-u* « ce cheval »). A noter que souvent (ouolof, sérère, diola, timné, etc.) le degré de détermination ou de proximité est indiqué par l'attribution de telle ou telle voyelle finale au déterminatif. Le nom de nombre suit en général le nom de la chose nombrée ; en ouolof, où il est un nom véritable ayant pour complément le nom de la chose nombrée, il le précède.

Les désinences consonantiques sont fréquentes, surtout dans les langues du Nord. Dans celles du Sud (timné, limba, boulom, etc.), on a tendance à n'employer que des désinences vocaliques.

Des tons musicaux à valeur étymologique existent dans les langues du Sud. D'autre part, le peul possède au moins un cas

de tons musicaux à valeur grammaticale : à l'aoriste présent et à l'aoriste d'habitude de la voix active, c'est le ton musical seul qui distingue le négatif de l'affirmatif.

Le pluriel est indiqué dans les noms par l'indice de classe, qui varie selon le nombre, et, éventuellement, par la modification de la consonne initiale de la racine (ouolof et limba quelquefois, sérère, peul). En boulom l'indice de classe, bien qu'ayant disparu le plus souvent au singulier, s'est maintenu au pluriel. Lorsque l'indice de classe est complètement tombé ou a été incorporé au radical (cas fréquent en ouolof), rien ne distingue le pluriel du singulier dans les noms indéterminés.

Voici, dans leur forme essentielle, les pronoms personnels de quelques langues.

	Peul	Ouolof	Sérère	Diola
1re pers. sing.	*mi*, *am*	*ma*	*mi*	*i*, *om*
— plur.	*min* (excl.), *en* (incl.)	*nu*	*in*	*di*
2e pers. sing.	*a*, *da*	*a*	*wo*	*u*
— plur.	*on*	*en*	*om*	*d'u*
3e pers. sing. commun	*dum*, *um*	*mu*	*a*, *le*	*mu*
— plur. —	*dumen*	*n'u*	*a*, *de*	*mu*

	Timné	Boulom	Limba	Kissi
1re pers. sing.	*i*, *mi*	*ya*, *mi*	*ya*, *am*	*i*, *mi*
— plur.	*sö*	*hö*	*min*	*ne*
2e pers. sing.	*mö*	*mu*	*ndo*	*nom*, *nu*
— plur.	*nö*	*nö*	*na*	*n'a*
3e pers. sing. commun	*ọ*	*wọ*	?	*o*
— plur. —	*ña*	*ña*	?	*a*

Suit la nomenclature des 24 langues du groupe :

1° le **peul** ou *pular* ou *fulfulde* ou *fulbere* ou *ful* ou *pul* ou *fula* ou *fulāni* ou *fellata* ou *fellāniya*, parlé dans le Foûta-Toro ou Foûta sénégalais, le Boundou et le Ferlo au Sénégal, dans le Foûta-Diallon en Guinée Française, dans le Massina au Soudan Français, dans le Liptako en Haute-Volta, dans l'Adamaoua en Nigéria et au Cameroun et dans un grand nombre de colonies de moindre importance disséminées çà et là en Afrique Occiden-

tale et jusqu'à l'Est du Tchad [1], par des pasteurs nomades ou semi-nomades de race blanche plus ou moins métissée, assez peu nombreux, dits plus proprement *Fulbe* (sing. *Pullo*) ou *Wororbe* ou *Wururbe* (sing. *Bororo*) ou encore *Pulli*, par de nombreux noirs agriculteurs et sédentaires tels que les « Toucouleurs » ou *Tokoror* ou *Tekārir* du Foûta sénégalais ou Tekrour et du Boundou, par des métis de Peuls et de nègres divers, tels que les *Fula* du Foûta-Diallon et bien des gens dits improprement *Fulbe* du Soudan Français, de la Nigeria, etc., enfin par quantité de serfs nègres dits *Rimāybe*, vassaux des diverses catégories précédentes (XIe siècle El Bekri, XVIIe siècle vocabulaire anonyme publié en 1845 par d'Avezac);

2° le **ouolof** (*wolof*) ou *volof* ou *g'olof* ou *d'olof*, parlé sur le bas Sénégal à partir et en aval de Dagana et au Sud jusqu'à Dakar inclus (XVIIe siècle vocabulaire susmentionné);

3° le **none** (*nōn*) ou *d'oba*, dit « sérère-none », dans la région de Thiès (1868 Mage);

4° le **sérère** (*serẹr*) ou *kegem* ou *nd'egem*, dit « sérère-sine », dans le Baol, le Saloum et le Sine, à l'Est et au Sud de Dakar (XVIIe siècle voc. anonyme susmentionné);

5° le **diola** (*d'ola*) ou *yola* (dialectes *flup* ou *felup* et *filham* ou *foñ'i*), sur la basse Gambie et la basse Casamance et entre les deux (XVIIe siècle voc. anonyme susmentionné);

6° à 12° sept langues parlées entre la Casamance et l'estuaire de Boulam : le **bagnoun** (*bañ'un*), entre Casamance et Cacheo (XVIIe siècle voc. anonyme susmentionné); — le **balante** (*balāt*) ou *bulanda*, à l'Est du précédent, entre Casamance et Geba ; le **mandjak** (*mādžak*) ou *mānd'ago* ou *kañ'op*, sur les deux rives du bas Cacheo ; le **bôla** (*bōla*) ou *bulam* ou *burāma*, sur la côte, entre le Cacheo et le Geba (dialecte *sarar* ou *sadal*) et dans l'île de Boulam (dialecte *burāma*) ; le **papel** ou *pẹpel*, rive droite de l'estuaire du Geba et île Bissao ; le **biafare** (*biafare*) ou *biafada* ou *fada*, rive gauche du même estuaire ; le **bidjougo** (*bidžugo*) ou *bidžogo* ou *bidžago* ou *bisago*, archipel des Bissagos (pour ces six langues, 1854 Koelle);

1. Les Peuls établis au Darfour auraient adopté la langue arabe.

13° à 16° quatre langues parlées de la haute Gambie au haut Rio Nuñez et englobées souvent sous le nom commun de *tenda* : le **bassari** (*basari*) ou *biyan* ou *ayan* ou *wo* (néant) ; — le **koniagui** (*kon'agi*) ou *awõhẽ* ou *azen* (1913 Migeod) ; — le **badiar** (*bad'ar*) ou *padžad* ou *padžãde* ou *bigola* ou *agola* ou *axus* (1854 Koelle) ; — sous réserves le **tiapi** (*t'api*) ou *t'apesi* (néant) ;

17° à 23° sept langues parlées le long ou près de la côte depuis l'estuaire de Boulam jusqu'à la rivière Soulimah sur les frontières du Sierra-Leone et du Libéria : le **nalou** (*nalu*), entre le Rio Grande et le Rio Nuñez (1854 Koelle) ; — le **baga**, sur la côte, du Rio Grande à Conakry (1849 Clarke sous le nom de *barka* et 1854 Koelle) ; — le **landouman** (*landumã*) ou *landọma*, entre le haut Rio Nuñez et le haut Rio Pongo, à l'Est du nalou et du baga, et le **limba** ou *limbã* (dialectes *sella*, *safroko*, etc.), entre le soussou et le loko (1854 Koelle) ; — le **timné** (*timne*) ou *temne* ou *temene* ou « timmanee », avec le dialecte *sanda*, vallées de la Kolente, de la Petite Scarcie et de la Roquelle, et le **boulom** (*bulom*) ou *bullom* ou *šerbro* ou *mampwa*, sur la côte, de Freetown à Sherbro inclus (1803 Winterbottom) ; — le **krim**, entre Sherbro et la rivière Soulimah (1916 Thomas) ;

24° le **kissi** (*kisi*) ou *gibi*, isolé au Nord du gola, entre le toma et le guerzé à l'Est, le kono et le gbandi à l'Ouest (1827-1828 Kilham).

BIBLIOGRAPHIE. — Aucun ouvrage d'ensemble. A citer : pour le peul, H. GADEN, *Le poular, dialecte peul du Foûta Sénégalais*, Paris, 1913-1914, 2 vol. in-8, et Sylvia LEITH-ROSS, *Fulani Grammar*, London, 1922, in-8 ; pour le ouolof, RAMBAUD, *La langue wolof*, Paris, 1903, pet. in-8 ; pour le sérère, H. GREFFIER, *Dictionnaire français-sérèr précédé d'un abrégé de la grammaire sérère*, Saint-Joseph de Ngasobil, 1901, in-32; pour le diola, Ed. WINTZ, *Dictionnaire français-dyola et dyola-français précédé d'un essai de grammaire*, Paris, 1909, in-8 ; pour le biafare, KRAUSE, *Die Fada-Sprache am Geba-Flusse* (*Zeit. für Afrikanische Sprachen*, 1895) ; pour le timné, SCHLENKER, *Grammar of the Temne language*, London, 1864, in-8 ; pour le boulom, A. T. SUMNER, *A handbook of the Sherbro language*, London, 1921, in-8.

Appendice A. — Parlers des négrilles du Soudan.

En divers points du Soudan, du haut Nil jusqu'au Cameroun, on rencontre, dispersés par petits groupements, vivant en général uniquement de la chasse, des hommes de petite taille dits « négrilles », qui, au point de vue anthropologique comme à celui des mœurs, se distinguent nettement des nègres et se rapprochent beaucoup des « Bushmen » de l'Afrique Australe. Les parlers de ces négrilles n'ont fait l'objet d'aucune étude approfondie. Les quelques petits vocabulaires et renseignements recueillis par de rares voyageurs tendraient à faire croire que chaque groupement de négrilles soudanais parle la langue du peuple nègre auprès duquel il vit ou avec lequel il a été en relations dans le passé. Il est possible cependant qu'entre eux ces gens parlent des idiomes distincts des langues négro-africaines et présentant des affinités avec les idiomes des Bushmen et des Hottentots. La question reste à élucider. Dès maintenant, il semble attesté que, même lorsqu'ils usent de langues négro-africaines, les négrilles y ont introduit des vocables étrangers, des tons musicaux spéciaux et variés et des sons rappelant les clics des Bushmen et des Hottentots.

INDICATIONS BIBLIOGRAPHIQUES. — On trouvera quelques mots d'un parler des négrilles dits *Efe* ou *Efifi* ou *Vot'su* ou *Tiketike* ou *Akka*, répandus à l'Ouest du haut Nil et sur le haut Ouellé, dans G. CASATI, *Dix années en Equatoria*, trad. L. de Hessem, Paris, 1892, in-4 (p. 113) ; des parlers des négrilles dits *Mpagga* ou *Mbakka* de l'Oubangui et de la Lobaye dans Dr OUZILLEAU, *Notes sur la langue des Pygmées de la Sanga* (*Revue d'ethnographie et de sociologie*, 1911, pp. 75-92, voc. 4, 5 et 6) ; des parlers des négrilles dits *Babiñga* ou *Bayaga* de la Sanga dans le mémoire précité du Dr OUZILLEAU (voc. 1, 2 et 3), dans G. BRUEL, *Les populations de la Moyenne-Sanga : les Babinga* (*ibid.*, 1910, pp. 111-126), et dans Drs POUTRIN et GRAVOT, *Travaux scientifiques de la mission Cottes au Sud-Cameroun*, Paris, 1911, in-8 (III, pp. 80-101).

Appendice B. — Parlers négro-européens de la Guinée.

A l'occasion de la traite des esclaves et à la suite des contacts qui en ont résulté entre les Européens et les Noirs, il s'est formé

sur la côte occidentale d'Afrique des parlers mixtes qui ont acquis depuis une extension de plus en plus grande et qui sont employés couramment aujourd'hui par les Européens et les indigènes dans leurs relations mutuelles et même souvent par les indigènes entre eux lorsqu'ils ont des langues maternelles différentes. Chacun de ces parlers a composé son vocabulaire à l'aide de mots, plus ou moins déformés phonétiquement, empruntés à celle des langues européennes qui lui a donné naissance, beaucoup de ces mots étant des termes spéciaux aux marins ou provenant des parlers créoles de l'Amérique ; on rencontre aussi des mots pris à d'autres langues, par exemple, dans le négro-anglais, des mots pris au portugais ou à l'espagnol, comme *save* « savoir ». Quant à la grammaire, elle est conforme aux principes grammaticaux de la langue négro-africaine dont les ressortissants ont le plus contribué à la formation et au développement du parler négro-européen considéré. Le verbe a été pris dans sa forme la plus simple ou la plus répandue, par exemple, dans le négro-français, au participe passé ou à l'impératif, et non point à l'infinitif, comme bien des gens le croient pour n'avoir considéré que des verbes appartenant en français à la 1re conjugaison, où l'infinitif sonne comme le participe passé : *vi* (pour « vu ») = voir, *fini* = finir, *pri* ou *prã* = prendre, etc. On le conjugue au moyen d'affixes dont chacun a une valeur conventionnelle ; ainsi le parfait prend en négro-français le préfixe *ya* (il y a) et en négro-anglais le préfixe *dõn* (*done*, fait). L'ordre des mots varie selon les régions, se conformant à peu près à l'ordre des mots suivi dans le groupe négro-africain où le parler mixte s'est introduit ; c'est ainsi qu'en négro-français on dit au Sénégal *fam mwa* (femme moi) et à la Côte d'Ivoire *mõ fam* (mon femme) ou *mwa fam* (moi femme) pour « ma femme ».

Actuellement, trois parlers négro-européens sont en usage en Afrique Occidentale : le **négro-portugais**, parlé aux Iles du Cap Vert, d'où il paraît avoir fait disparaître l'ancienne langue indigène, en Guinée Portugaise, sur la Côte des Esclaves, où un nouvel apport lui est venu d'esclaves originaires de la région, transportés au Brésil, puis revenus dans leur patrie, et enfin dans

l'île de Sao-Thomé, où il a remplacé, semble-t-il, l'ancienne langue indigène ; — le **négro-anglais** ou « pidgin » (*pidzin*), répandu sur la basse Gambie, au Sierra-Leone, au Libéria et sur toute la côte depuis le Cap des Palmes jusqu'à l'équateur, ainsi que dans les centres, à l'intérieur des colonies britanniques ; — le **négro-français** ou « petit-nègre », parlé seulement dans les colonies françaises.

INDICATIONS BIBLIOGRAPHIQUES. — Des études ont été publiées sur ces divers parlers. On pourra consulter : pour le négro-portugais, F. DE BARROS, *Lingua creola da Guiné portuguesa e do archipelago de Cabo Verde* (*Revista de estudas livres*, Lisboa, 1885-1886) ; PAULA e BRITO, *Dialectos crioulos portuguezes* (*Boletim da Soc. de Geogr.*, Lisboa, 1887); ALMADA NEGREIROS, *O dialecto de S. Thomé* (pp. 303-369 de *Historia ethnographica da ilha de S. Thomé*, Lisboa, 1895, in-8) ; H. SCHUCHARDT, *Kreolische Studien*, I-IX, Wien, 1882-1891 ; — pour le négro-anglais, K. LENTZNER, *Colonial English*, London, 1891, in-8 ; H. SCHUCHARDT, *Beiträge zur Kenntnis des Englischen Kreolisch* (*Englische Studien*, XII, 3) ; note dans le chapitre VIII de DELAFOSSE, *Vocabulaires comparatifs*, Paris, 1904, in-8 ; — pour le négro-français : L. ADAM, *Les idiomes négro-aryen et maléo-aryen*, Paris, 1883, in-8 ; H. SCHUCHARDT, *Beiträge zur Kenntnis des Kreolischen Romanisch* (*Zeit. für Romanische Philologie*, XII) ; note dans le chapitre VIII de DELAFOSSE, ouvrage précité.

Appendice C. — Parlers négro-africains de l'Amérique.

Dans plusieurs îles des Antilles (Haïti notamment) et dans diverses provinces de la Guyane et du Brésil, peut-être en d'autres régions de l'Amérique, les parlers négro-africains des esclaves amenés d'Afrique aux siècles passés ont subsisté jusqu'à ce jour parmi les descendants de ces esclaves, et sont encore employés, au moins au cours de certaines cérémonies magico-religieuses. Nous manquons de documentation quant à l'extension et à l'emploi de ces parlers, ainsi qu'en ce qui concerne les langues négro-africaines qu'ils représentent ou dont ils dérivent. Il semble cependant que le groupe éburnéo-dahoméen soit celui qui a fourni l'appoint le plus considérable et le plus persistant.

Maurice DELAFOSSE.

LES LANGUES BANTOU

GÉNÉRALITÉS

L'hypothèse de l'unité primitive des langues parlées à Zanzibar, dans l'Angola et dans la colonie du Cap a été émise dès 1808 par Lichtenstein. Vivien de Saint-Martin démontra l'unité non seulement de ces langues, mais aussi de celles du Congo dans une communication faite à la Société de Géographie de Paris en 1847; Ewald soutint la même thèse dans une communication publiée en 1848 dans la *Zeitschrift der deutschen morgenländischen Gesellschaft*, sans faire allusion au mémoire de Vivien de Saint-Martin qu'il doit avoir connu cependant. En 1856 William Bleek, bibliothécaire du gouverneur de la colonie du Cap, proposa de donner à ces langues le nom de « bantou ». Le vocable *bantu* signifie « les hommes » dans tous les idiomes du groupe ; sa forme simple et claire l'a vite fait adopter par tous les savants [1].

Le groupe bantou comprend la presque totalité des idiomes parlés dans le triangle qui a pour sommet le cap de Bonne Espérance, pour côtés : la côte orientale jusqu'au pays des Somali et la côte occidentale jusqu'à Douala, et pour base une ligne qui partant du pays des Somali passe au Nord de l'Ouganda, se rapproche du Congo et aboutit à l'Atlantique (voir la carte, page 589).

Il semble préférable, en l'état actuel de nos connaissances, de ne pas chercher à fixer le tracé exact des limites septentrionales du groupe.

Jusqu'ici toutes les langues du Gabon et une partie de celles

1. Le terme « bantou » est ici invariable, conformément aux règles adoptées pour l'ensemble du livre, mais contrairement aux habitudes de l'auteur du chapitre.

du Cameroun ont été considérées comme appartenant au groupe bantou, les idiomes assez similaires parlés au delà du Cameroun étant appelés semi-bantou. Il est plus correct de ne plus faire entrer le pahouin, le boulou etc., dans le groupe des langues bantou, dont ils se séparent sur plus d'un point, et d'envisager plutôt un groupe ouest-africain, attestant une langue commune sœur du bantou commun. L'unité des langues africaines étant généralement admise aujourd'hui et l'époque de communauté étant inconnue, le fait d'admettre deux groupes permet de faire un tableau plus net, mais ne signifie pas la négation a priori d'une période de communauté de ces deux groupes postérieure à l'unité africaine.

Dans le vaste triangle décrit ci-dessus, les seuls parlers qui ne soient pas bantou sont ceux des peuplades naines que l'on rencontre en certains points des forêts du Congo et de l'Afrique australe et la langue des Hottentots, peuple assez important, établi dans l'Afrique australe, à l'Ouest du désert de Kalahari.

Les langues bantou ont un vocabulaire et un système morphologique communs. Les différences dialectales sont dues à des évolutions phonétiques propres, bien que soumises à des lois générales communes. Ces évolutions ont eu pour corollaire des différences syntaxiques et sémantiques. Elles semblent plutôt attester l'exagération de certaines tendances individuelles que des influences étrangères.

Si l'unité linguistique est indiscutable, l'unité ethnique est moins nettement établie, et en l'absence de toute documentation ancienne, il est difficile de rattacher les Bantou à un groupe ethnique connu. Les traditions indigènes semblent indiquer que les courants d'émigration sont partis du haut Nil, mais on ne peut rien affirmer. Schleicher et Reinisch ont du reste proposé de relier le bantou au groupe sémitique.

Les Arabes ont été en contact avec des Bantou à Zanzibar depuis des siècles et le swahili, dialecte de ce pays, a été écrit en premier lieu en caractères arabes. Les autres langues bantou ont été notées par des Européens, qui ont employé les caractères latins modifiés, le cas échéant, par des points diacritiques.

Les auteurs arabes du moyen âge citent quelques mots bantou. A ceci près, ces langues ne nous sont connues que par les documents et les lettres de missionnaires, de voyageurs et de fonctionnaires européens.

Pigafetta donne quelques mots congolais dès 1591 et des jésuites se rendant aux Indes au XVIe s. donnent des renseignements sur une langue de la côte orientale dans leurs lettres à leurs supérieurs. Brusciotto publia une grammaire congolaise en 1650 ; un catéchisme avait été imprimé en 1624. Il y a remarquablement peu de différence entre les parlers modernes et ceux du XVIIe siècle. Nous savons donc que les langues ont peu varié au cours des quatre derniers siècles et que la répartition était sensiblement la même qu'aujourd'hui (au moins sur la côte).

Tout porte à croire que les langues des négrilles et des Hottentots ont reculé devant le bantou ; mais il est évident que dans les pays où les colons blancs sont nombreux comme au Cap, les idiomes noirs ont été très influencés, et ont perdu du terrain. Les limites septentrionales n'étant pas faciles à déterminer pour ce qui concerne les parlers modernes, il est impossible de dire si en ces régions il y a eu extension ou diminution du domaine bantou au cours des siècles.

Les statistiques anglaises donnent le nombre de 50 millions de nègres bantou ; ce chiffre paraît très exagéré.

DESCRIPTION PHONÉTIQUE [1]

Les langues bantou sont sonores et l'articulation y est généralement nette.

Toutes les syllabes se terminent par une voyelle et les seuls groupes de consonnes possibles sont *nasale-consonne* ou *consonne -w* (ou *y*). Exemples : *ku-bo-na*, *ba-ntu*, *ku-gwa*, *ku-lya*.

Les mots étrangers à groupes de consonnes sont modifiés profondément Ex. *Kirisiti* — Christ, *silivera* — angl. silver.

1. Lorsqu'un exemple est commun à la majorité des parlers, il est cité sans indication de langue.

Accent. — Dans tout mot de deux ou plusieurs syllabes l'avant-dernière est accentuée ; si l'on ajoute un suffixe l'accent se déplace. Ex. : *ku-tánda* : aimer, *ku-tandána* : s'aimer l'un l'autre.

Voyelles. — Le timbre ou degré d'ouverture des voyelles varie à l'infini. Certains auteurs ont essayé de noter des timbres multiples. Aucun système de notation par les lettres ne permet de noter toutes les modulations de la voix humaine ; le timbre varie d'individu à individu et de génération à génération, et aussi en fonction de la longueur du mot et de la nature des phonèmes environnants. On trouve dans toutes les langues bantou *a*, *e* et *o* ouverts, *i* et *u*. En général, on trouve aussi *e* et *o* fermés ; en certains dialectes, le degré de fermeture est tel qu'il est difficile de savoir si l'on entend *e*, *o* ou *i*, *u*.

Les voyelles mixtes sont rares ; on a signalé *ü* en herero et dans la région du Kilimandjaro.

Les voyelles longues représentent généralement sinon toujours la contraction de deux voyelles à la suite de l'amuïssement d'une consonne. Ex. : douala *mā* ≐ *maka*.

Ton. — Le ton musical ou modification de la hauteur d'émission des voyelles a été signalé en quelques groupes, en nyamwézi par exemple, mais il ne semble pas avoir l'importance sémantique qu'il a dans les idiomes ouest-africains, soudanais, etc.

Consonnes. — Dans la presque totalité des idiomes on rencontre trois consonnes occlusives sourdes *p*, *t*, *k* [1], trois sonores *b*, *d*, *g*, quatre nasales *m*, *n*, *ñ*, *ṅ*, les consonnes *l* (ou *r*) *w*, *y* et quatre affriquées et fricatives qui sont soit les chuintantes *tš*, *š*, *dž*, *ž*, soit les sifflantes *ts*, *s*, *dz*, *z*. Les phonèmes *k′ g′*, *x*, *g*, *ƥ*, *đ*, *b*, *f*, *v* se rencontrent dialectalement. Dans le groupe du Sud-Est (cafre-sotho) on rencontre des clicks ou sons à inspiration buccale qui ont été vraisemblablement empruntés aux Hottentots ou aux Bochimans. Les laryngales (sauf *h*) sont inconnues.

Alternances consonantiques. — Dans toutes les langues bantou il y a des alternances consonantiques qui sont régulières. Ainsi

1. Ces occlusives présentent une variante, dont la nature n'est pas bien déterminée et qui a été notée ici par la consonne suivie de ^{h}, ainsi t^h.

les mots qui commencent par une mi-occlusive après une voyelle ont une occlusive initiale après une nasale préfixe. Ex. :

yao *a m-bonile* « il m'a vu » *a tu-wonile* « il nous a vus ».
a n-dolite « il m'a regardé » *a tu lolite* « il nous a regardés ».
ganda *ndere* « lacets » *olulere* « un lacet ».
soubiya *inombe ondotu* « beau bœuf » *luzim bolulotu* « belle chanson ».

Quelquefois la nasale s'est amuïe, mais les rapprochements interdialectaux montrent qu'il s'agit d'une ancienne nasale ou bien d'un élément fermé (consonne ou *i*). Ex. :

ganda	*edobo*	« crochet »	pl. *amalobo*.	
galoa	*ibega*	« épaule »	*owega*	
sotho	*ho it^hata*	« s'aimer »	*ho rata*	« aimer ».
	ho ik^hanèla	« se refuser »	*ho hanèla*	« refuser ».

En beaucoup de dialectes, la consonne finale du radical varie selon que la voyelle finale suffixée est *-a*, *-o*, ou *-i*, *-u*. Ex. :

ganda	*ku longoka*	« être propre »	*longofu*	« propre ».
	ku longa	« ensorceler »	*lovu*	« ensorcelé ».
	gonda	« être doux »	*gomvu*	« doux ».
yao	*langa*	« dire adieu »	prétérit	*lanžile*.
	tenda	« faire »	«	*tesile*.
	puga	« souffler »	«	*pužile*.
soubiya	*subila*	« rougir »	«	*subidi*.

Restitution du système phonétique du bantou commun.

Les correspondances phonétiques montrent : 1° qu'il a dû y avoir en bantou commun quatre occlusives sourdes **p* **t* **k'* **k*, quatre sonores **b*, **d*, **g'*, **g*, quatre nasales **m*, **n*, **ñ*, **ṅ*; 2° qu'il y a lieu de distinguer au moins sept voyelles **a*, **ǫ*, **ę*, **ọ*, **ẹ*, **i*, **u* ; les deux dernières attestant vraisemblablement des diphtongues.

Toute l'évolution paraît avoir été soumise aux lois suivantes :

1° Toute consonne tend à assimiler son point d'occlusion à celui des phonèmes en contact.

2° L'un des phonèmes en contact étant fermé, l'occlusion des sonores a été maintenue ; autrement l'occlusive est affaiblie dialec-

talement ; et l'affaiblissement a été en rapport direct avec le degré d'ouverture des voyelles. Ex. :

Sotho *laya* = thonga *laya* « ordonner ».
lema = *rima* « cultiver ». (*r* est plus fermé que *l*).

Les sourdes présentent une stabilité remarquable ; toutefois dialectalement certaines ne sont plus représentées que par un souffle en position intervocalique ou à l'initiale. Ex. :

Sotho	*ho-fa*	« donner »	nyoro	*ku-ha*.
	ho-p^heta	« idée de courber »	»	*ku-henda*.
	ho-ip^homola	« se reposer »	»	*ku-humura*.
	ho-hama	« traire »	»	*ku-kama*
	ho-hola	« grandir »	»	*ku-kura*

Remarquons enfin que *r* alterne tantôt avec *d* tantôt avec *t* et doit donc être considéré comme une mi-occlusive dentale. Ex. :

Sotho *ho-rata* « aimer » *ho-it^hata* « s'aimer ».
Galoa *tonda* « aimer » *ronda* « aime ».

MORPHOLOGIE

Les mots bantou se composent d'un radical et d'affixes suffixés ou préfixés dont la valeur grammaticale est très nette et qui se retrouvent dans tous les parlers du groupe.

Le plus grand nombre des radicaux simples sont de la forme *consonne-voyelle-consonne* ou bien *consonne-voyelle-nasale-consonne*. Les mots à voyelle initiale sont rares ; les radicaux monosyllabiques terminés par une voyelle sont peu nombreux, mais il en existe un certain nombre communs à tout le groupe. Ex. :

Douala	*lom-*	« idée d'envoyer »	Sotho	*rom-*.
«	*tond-*	« idée d'aimer »	«	*rat-*.
«	*alab-*	« idée de répandre »	«	*arab-*.
«	*be*	« être, devenir »	«	*ba*.

Les radicaux de la forme *consonne*-w-*voyelle-consonne*, sont communs à un très grand nombre de dialectes, mais ne représentent pas des radicaux simples primitifs. Ex. :

Cafre, yao, bisa *kwela* « idée de monter ».

Les suffixes modifient la valeur du radical, tandis que les préfixes indiquent le mode d'emploi et les rapports.

Ainsi les préfixes *ku*, *u* caractérisent le premier l'infinitif, et le second la 2e pers. sing. de l'aoriste, tandis que le suffixe *-a* indique la durée ; on aura donc *ku-tand-a* « aimer », *u-tand-a* « tu aimais et tu aimes » (action prolongée).

Le suffixe *-i* indique qu'il s'agit d'une action achevée ou d'un état acquis ; employé conjointement avec le préfixe *mo-* ou *mu-* caractéristique des noms de personnes, ce suffixe forme les noms d'agents. Ex. :

Sotho *ho laela* « commander » *mo-laeli* « commandeur ».
Nganja *ku sula* « forger » *m-suzi* « forgeron »
(*l* devant *a* alterne avec *z* devant *i*).

Les suffixes suivants se retrouvent dans tous les parlers avec une valeur sensiblement semblable.

-a idée de durée, d'indétermination ; suffixe de l'infinitif, de l'aoriste, de l'impératif, et d'un grand nombre de substantifs. Ex. : Sotho *lala* « passer la nuit, » *molaetsa* « message, commission ». Sotho, Thonga, etc. *mala* « entrailles ».

-e suffixe de quelques adjectifs adverbiaux, et d'un grand nombre de substantifs ; employé dans les formes verbales négatives ou dubitatives, ce suffixe représente vraisemblablement **a* + **i* négatif. Ex. : *kale* « ancien ».

-i idée de l'action achevée, d'un état acquis, de quelque chose de déterminé ; forme les thèmes des noms d'agent (voir ci-dessus) et apparaît dans le suffixe du prétérit (voir ci-dessous).

-o suffixe nominal, forme le nom verbal et un grand nombre de substantifs. Ex. :

Sotho *ho-it^{h}ata* « s'aimer » *t^{h}ato* « amour »
mokuti « coiffeur » *mokuto* « action de coiffer ».
Ganda *lagira* « commande » *kilagiro* « ordre ».

-u suffixe qualificatif (lorsqu'il ne s'agit pas de $u < {}^*\varrho$), donne des adjectifs (considérés quelquefois comme des participes) et quelques substantifs de qualité. Ex. :

Sotho *sefofu* swahili *mpofu*, « aveugle ».
Ganda *longofu* « propre ».

-ile (dialectalement *ire*, *-ide* etc.) suffixe du prétérit dans tout le groupe et quelquefois du parfait. Ex. *o bonile* « il a vu ».

-ela suffixe du thème verbal directif ou déterminé, c'est-à-dire du thème employé lorsque le but de l'action ou de l'état est exprimé ou lorsqu'il s'agit d'une action ou d'un état unique déterminé. Ex. :

Sotho *loka* « être bien, droit, juste ».
lokela « convenir à, être bon pour ».
lula « s'asseoir ».
lulela « s'asseoir pour attendre quelqu'un ».

-isa suffixe du thème verbal causatif. Ex. :

Sotho *lula* « s'asseoir ».
lulisa « faire asseoir ».

-ana suffixe du thème verbal réciproque. Ex. :

Sotho *opa* « frapper ».
opana « se frapper (l'un l'autre »).

-eka (dialectalement *-ika*, *-eha*) suffixe du thème verbal statif : le sujet est considéré comme étant dans un état ou comme en train de faire l'action ou bien comme susceptible de la subir. Ex. :

Sotho *ho hlomp^ha* « honorer » *hlomp^heha* « être honorable ».
Ganda *ku laba* « trouver » *ku labika* « être trouvé ».

-ola (dialectalement *-ula*, *ura*, *ua*).
-osa (» *usa*)
-oka (» *uka*, *-oha*)

sont des suffixes qui donnent des thèmes inversifs correspondants à ceux formés avec les suffixes *-ela*, *-isa*, *-eka*. En réalité c'est la voyelle qui est vraiment l'élément d'inversion. Le plus souvent on pourrait qualifier ces suffixes de négatifs ; mais un examen attentif montre qu'ils indiquent plutôt le sens opposé que la négation. Ex. :

Sotho *ola* « ôter les gros morceaux à la main avant de vanner ».
olosa « vanner (en agitant le grain sur un plateau »).
hata « fouler, piétiner ».
hatola « cesser de piétiner ».
roala « poser une charge sur la tête ».
rōla (= *roola*) « déposer une charge ».

Ces suffixes n'ayant pas encore été l'objet d'une étude approfondie, il n'est pas possible d'en rétablir ici les formes primitives; toutefois il est difficile de ne pas être frappé par le rôle particulièrement important joué par les voyelles, rôle sur lequel on a peu insisté jusqu'ici. Ce sont elles qui modifient et déterminent la signification du groupe consonantique qui constitue le mot; les affixes consonnes ne font que développer le sens du radical primitif. Ex. :

Sotho *o bonile* « il a vu », *o bonela* « il pourvoit à »
ho finella « nouer fortement »; *ho finolla* « dénouer ».

Il est donc vraisemblable qu'en africain commun le rôle morphologique respectif des consonnes et des voyelles était analogue à celui de ces éléments dans les langues sémitiques.

Il nous reste à exposer le système des préfixes dont le rôle est si remarquable que les idiomes bantou sont presque toujours appelés langues à préfixation.

On peut réunir les morphèmes préfixés en trois groupes.

A, pronoms personnels ; B, préfixes à nasale ; C, préfixes sans nasale.

A. Les préfixes pronoms personnels de la 1re et 2e personne singulier et pluriel s'emploient comme sujets à l'initiale d'une forme verbale et comme compléments devant le verbe.

1re pers. sing. **me* ou **ni*, pl. **to* ou *tu*.
2e » » **go* ou **gu*, » *mu* ou *nu*
» » **ko* ou *ku* (Compl. dir.).

B. Les préfixes à nasale sont caractéristiques des noms. Ce sont :

mo (dialectalement *mu*, *m*)
me (» *mi*)
ma
n (*m* devant une labiale)

Ex. : *mu-ntu* « homme ». — *mu-kazi* « femme ».
ma-la « entrailles ».
mi-ti « arbres ».
n-koko « poule, volaille ».

C. Les préfixes sans nasale sont les pronoms de la 3e personne ;

quelques-uns s'emploient devant un thème pour en préciser la valeur nominale ou pour le qualifier. Ex. : *le bona*, il voit, *le-bone* « bougie ».

En voici les formes hypothétiques en bantou commun avec un aperçu des formes modernes les plus fréquentes.

**g'o*	(*o, u, yu*)
**go*	(*o, u, gu*)
**g'e*	(*e, yi, ẓi*)
**g'a*	(*a, ya, ga*)
**di*	(*li, ẓi, zi*)
**de*	(*le, li, ri*)
**do*	(*lo, lu, ru,*)
**ke*	(*ki, se, tši*)
**ko*	(*ku, ho*)
**ka*	(*ha*)
**ba*	(*ba, wa, va, a*)
**bi*	(*bi, vi*)
**bo*	(*bo, bu, wu, vu*)
**tu*	(*tu*)

Aujourd'hui tous les substantifs bantou se groupent en classes, chaque classe étant caractérisée par l'emploi de tel ou tel préfixe devant tous les morphèmes ou mots se rapportant à un substantif. Ex. :

Sotho	*marapo a* (**g'a*) *mane*	« quatre courroies ».
	batho ba bane	« quatre hommes ».
Ganda	*ekibya kino ekibi*	« cette mauvaise terrine ».
pl.	*ebibya bino ebibi.*	
	ekibia kiri ekinene	« cette grande terrine là-bas ».
	ebibya bya tulaba	« les terrines que nous voyons ».

Les bantouistes ont toujours affirmé jusqu'ici que tous les préfixes correspondent à des conceptions matérielles, et les hypothèses les plus diverses ont été formulées sur leur origine. En l'état actuel de nos connaissances, nous nous bornerons à indiquer les faits communs à la majorité des parlers.

Les préfixes à nasale indiquent toujours un substantif. En

quelques dialectes un adjectif se rapportant à un substantif à préfixe nasal prend le même préfixe, mais le préfixe nasal n'est jamais employé comme pronom. Ce sont les pronoms supposant *g'o, *go, *g'e, *g'a, qui s'emploient comme relatifs, comme démonstratifs et comme pronoms de la 3ᵉ personne avec les substantifs à préfixe nasal. Les autres pronoms du groupe C sont tantôt préfixés à un thème employé nominalement tantôt employés comme pronoms. Voici quelques exemples de l'emploi des préfixes et des indications sur les faits exprimés par leur opposition.

Le préfixe *mo* ou *mu*, lorsque le préfixe pronominal correspondant repose sur *g'o*, est le morphème particulier du singulier des noms désignant des êtres humains. Le préfixe correspondant du pluriel est *ba*, etc. Ex. :

Ganda, etc.	*muntu*	« homme »	plur.	*bantu*.
	mulogo	« sorcier »	»	*balogo*.
Sotho	*mosotʰo*	« un homme sotho »	»	*basotʰo*.
	mofumi	« un richard »	»	*bafumi*.

Mo ou *mu*, lorsque le pronom correspondant repose sur **go* est caractéristique du singulier de noms désignant des objets très variés et quelquefois des personnes assimilées à des choses. Le pluriel de ce nom est caractérisé par le préfixe nominal *me* ou *mi* et par un pronom attestant **g'e*. Ex. :

Ganda, etc.	*muliro*	« feu ».		
Sotho, »	*mono*	« doigt »,	plur.	*meno*.
Nyamwezi	*muti*	« arbre »,	»	*miti*.

Ma est caractéristique des noms désignant un ensemble, une collectivité, un très grand nombre ; le pronom correspondant suppose **g'a*. Lorsqu'on veut parler d'un seul élément de la collectivité on emploie un préfixe qui remonte à **de*. Ex. :

Sotho	*manaka* « les cornes (d'un animal) »,	*lenaka* « une corne ».
Bisa	*matama* « les joues »	*litama* « une joue ».

Dialectalement on trouve des noms d'ensemble avec préfixe *li* etc. (de **de*) Ex. nyamwezi *ivu* « cendres ».

N est par excellence le morphème caractéristique du nom. Il se trouve devant la plupart des noms d'animaux et devant les noms dérivés de verbe. Le pronom correspondant atteste **g'e* au

sing. et *di au pluriel. Au pluriel, ce dernier pronom est souvent préfixé à la nasale initiale mais *ne la remplace pas*. Ex. :

nyota « étoile » pl. *dinyota*.
mbuzi « chèvre » » *mbuzi*.
Sotho *tlou* « éléphant » » *litlou*.

Il existe un certain nombre de noms désignant surtout des choses plurielles à préfixe attestant **din* qui prennent un préfixe attestant **do*, lorsqu'on veut parler d'une unité. Ex. : nyamwezi *nzwili* « cheveux », *luzwli* « un cheveu ».

Les préfixes attestant **ke*, pl. **bi* sont surtout employés lorsqu'il s'agit d'objets fabriqués. Dialectalement ils désignent les choses propres à un pays et exceptionnellement des personnes assimilées à des objets. Ex. :

Sotho : *mosot^ho* « un homme sotho » ; *sesot^ho* « chose sotho, la langue sotho » *sep^heli* « un vivant ».

Soubiya *ši-ntu* « chose », pl. *bintu*.
Nyamwezi *ki-neneko* « gobelet », pl. *fineneko*.

Le préfixe *ku* etc., attestant **ko*, est caractéristique du nom verbal ou infinitif dans presque tout le groupe. Il est quelquefois locatif et avait vraisemblablement en bantou commun la même valeur que les prépositions *à* en français et *to* en anglais. Ex :

Ganda *ku-lagira* « commander »,
Sotho *ho-bona* « voir ».

Les préfixes *ka* et *tu* sont diminutifs. Ex. :

muntu « homme », pl. *bantu*.
kantu « petit homme », pl. *tuntu*.

Le préfixe attestant **bo* est caractéristique des noms abstraits mais apparaît en deux ou trois noms concrets communs à tout le groupe. Ex. :

bubi « le mal » (*-bi* « mauvais »).
vugeni « voyage » (**genda* « voyager »).
bwato « canot ».

Dans les idiomes du centre les prépositions locatives attestant **mo*, **pa* sont traitées comme de véritables préfixes d'accord, et dialectalement on cite des préfixes augmentatifs, péjoratifs, etc.

inconnus à la majorité des parlers ; mais ces faits isolés peuvent attester aussi bien un développement particulier du système que des survivances, et ne sont pas assez importants pour figurer dans un tableau d'ensemble.

De la distinction des catégories grammaticales.

Tous les affixes que nous venons d'énumérer ont certainement une valeur nette (bien que pour certains celle-ci nous échappe encore quelquefois), mais les catégories grammaticales africaines ne répondent pas à celles des langues latines.

On peut rendre toutes les idées exprimées par nos modes et temps en recourant à des particules et à des verbes auxiliaires, mais il ne faut pas identifier les formes correspondantes. L'africain distingue en effet la description de l'action et la description de l'état du sujet agissant. Dans le premier cas, il emploie une forme verbale pure, dans le deuxième il préfère souvent des formes qui correspondent vraiment à des prépositions attributives. Exemples :

Nyarwanda *ba mu bwira* « on lui dit ».
mmaze kupfa « je serai mort » (m. à m. je serai fini de mourir).
Xosa *u ya lila* « il pleure » (m. a. m., il est pleurant).
u lila « il pleurait ».

La différenciation entre le verbe et le substantif est donc assez délicate à définir, car les grammairiens européens considèrent comme verbales des formes qui sont vraiment nominales. Afin de ne pas soulever des questions dont la discussion dépasserait le cadre de cet exposé, nous nous bornerons aux faits indiscutables.

Les préfixes à nasale et les suffixes *o*, *u*, sont des indices d'une forme nominale dans la presque totalité des parlers.

Les langues qui, comme le herero, peuvent avoir *o*, *u* par assimilation à la finale d'une forme verbale, sont très peu nombreuses.

Par contre les pronoms personnels sujets et les préfixes unique-

ment pronominaux (attestant *g'o, *g'e, *g'a) indiquent toujours une forme verbale. Ex. :

Sotho *mp^{h}o* « don », *ofa*, « il a donné ».
molaeli « chef, commandant » *re laela* « nous ordonnons ».
Ganda *omutego* « piège » *ku tega* « poser un piège ».
tegevu « intelligent » *tu tegera* « nous comprenons ».

Quant aux mots à voyelle finale *e* ou *a* ou à préfixe sans nasale, rien ne permet a priori d'en distinguer la valeur, et la syntaxe seule montre s'il faut les traduire par un substantif ou par un verbe.

Un thème seul sans préfixe et avec un suffixe *-a* est nécessairement un verbe à l'impératif affirmatif. Ex :

lya « mange ».
mala « finis ».

Le pronom sujet est à l'initiale de toute forme verbale, même lorsque le substantif est à l'initiale de la proposition, et le pronom complément est préfixé au verbe.

Le substantif sujet précède toujours le verbe, et le substantif complément est postposé; en certains dialectes le pronom complément de même classe est préfixé au verbe même lorsque le substantif complément le suit.

Nyarwanda *Nyiragahinda a ra ba menya* « Nyiragahinda les connait ».
Tonga *Balavu babaluma bantu* « les lions mordirent les hommes ».
Ba lo mu tenda « ils le brûlent ».

En général lorsque le complément limite l'action ou la détermine, le verbe est à la forme déterminée; lorsque le complément est locatif ou indirect, il est séparé du verbe par des prépositions qui ont quelquefois été assimilées aux préfixes classificateurs. Ex. :

Tonga *Bamue ba kede ku Bulumbu* « les uns demeurent à Bouloumbou ».

Ganda *ya leta* « il apporta », *ya mu leterà* « il lui apporta de la nourriture ».

Le complément du nom (substantif ou pronom) est postposé

et rattaché au substantif déterminé par une particule attestant **g'a* précédée elle-même du pronom préfixe de ce substantif. Ex. :
Ganda *omuti gwa kabaka* « à l'arbre du roi »,
pl. *emiti gya kabaka.*
omuti gwa nge « mon arbre » (m. à m. arbre de moi),
pl. *emiti gya nge.*

Pour compléter cet aperçu sommaire du système morphologique bantou il y a lieu de signaler :

1° la rareté des qualificatifs, la plupart des adjectifs européens se traduisant par des verbes de qualité. Ex :
Tonga *muntu ua ka tuba ku mutue* « homme à cheveux blancs » (m. à m., « qui a blanchi sur la tête »).
Nyarwanda *umuntu w'omutima* « un homme courageux » (m. à m. un homme de cœur).

2° l'opposition généralement très nette de formes verbales positives et négatives. Ex. :
Soukouma *a li tula* « il frappe » *a tu tula* « il ne frappe pas ».
a turire « il a frappé » *a ta turire* « il ne frappa pas ».

3° l'existence de la numération décimale dans la plupart des parlers ; mais la diversité des formes laisse supposer qu'elle ne remonte pas au bantou commun.

CLASSIFICATION DES LANGUES BANTOU.

On trouvera ci-après une énumération des parlers bantou qui ont fait l'objet de travaux importants. Chaque village ayant quelques traits phonétiques, voire même quelques tournures, en propre, certains auteurs ont dressé de longues listes de parlers; mais il ne semble pas utile de faire de cette liste un répertoire géographique. Pour les groupes importants eux-mêmes, les caractéristiques sont peu saillantes, et la répartition des dialectes est souvent arbitraire faute de distinctions nettes.

Les langues du groupe sud-est présentent des clicks empruntés aux Hottentots; elles se divisent en deux sous-groupes : xosa-zoulou et sotho-tchouana, ce dernier caractérisé par la disparition

dn groupe phonétique *nasale-consonne sonore*, remplacé par des sourdes douces. Ex. :

Xosa *itanda* « s'aimer », sotho *it^hata*.
ku bombà « plâtrer », *ho bopa*.

Les langues du groupe sud-ouest (herero, etc.) présentent de nombreux faits d'assimilation de voyelles.

Les langues du groupe nord-ouest (galoa, douala, etc.) ont un système de classes moins net et plus simple, tandis que celles du groupe ganda présentent des préfixes inconnus ailleurs.

Il est probable que notre exposition déroutera les bantouistes habitués à certains clichés. Si nous nous sommes écartée quelque peu des conceptions de Bleek admises jusqu'ici par tous les auteurs, c'est parce que nous estimons qu'il a décrit admirablement les faits modernes, mais que ceux-ci attestent des développements parallèles et non pas l'état ancien.

L'unité primitive des langues parlées par les noirs d'Afrique paraît certaine, et on comprendra les faits bantou lorsque l'on connaîtra mieux l'africain commun.

BIBLIOGRAPHIE

DES PRINCIPAUX TRAVAUX D'ENSEMBLE SUR LES LANGUES BANTOU

Note. — Les principales abréviations employées dans cette bibliographie et les suivantes sont :

ZDMG. Zeitschrift der deutschen morgenländischen Gesellschaft.
WZKM. Wiener Zeitschrift für die Kunde des Morgenländes.
MSOS. Mittheilungen des Seminars für Orientalische Sprachen (Berlin).
JRAS. Journal of the Royal Asiatic Society.
ZAOS. Zeitschrift fur Afrikanische und Ozeanishe Sprachen.
JAS. Journal of the African Society.

AVELOT (Cap.), *Migrations dans l'Ogooué*. Bull. géogr., hist. et descript., 1906.
BASSET (R.), *Rapport sur les langues africaines*, XIe cong. des Orientalistes.
BLEEK (W. H.), *Languages of Mozambique*. Londres, 1856.
— *Comparative handbook of Australian African and Polynesian languages*, 1858.
— *Catalogue of Sir G. Grey's library Capetown*, 1858.
— *Comparative grammar of South African languages*. Londres, 1869 [1].

CLARKE (J.), *Specimens of dialects.*

CUST (R. N.), *Modern languages of Africa*, 1883 (trad. franç.)[2].

EWALD (H. v.), *Über die Völker und Sprachen südlich von Aethiopien.* ZDMG., 1847.

FINCK, *Verwantschaftsverhältnisse der Bantusprachen.* Göttingue, 1908 [3].

GREGORIO (G. di), *Cenni di glottologia Bantu.* Turin, 1882.

— *Origine significative dei cosidetti prefessi.* St. Glott. It., 1907.

HOMBURGER (L.)., *Phonétique historique du bantou.* Paris, 1913.

JACOTTET, *Bantu phonetics.* Lovedale, 1907.

JOHNSTON (Sir H.), *A comparative study of the Bantu and semi-Bantou languages.* Londres, 1919 [4].

— *Basis for a comp. gram. of the Bantu lang.* JRAS., 1907.

— *The Uganda Protectorate.* Londres, 1902.

KOELLE (S.), *Polyglotta africana.* Londres, 1854 [5].

KOLBE (F. W.), *Language study based on Bantu.* Londres, 1888.

LAST (J. F.), *Polyglotta africana orientalis.* Londres, 1885 [5].

LEPSIUS (C. R.), *Nubische Grammatik*, 1880 [6].

LICHTENSTEIN, *Mithridates*, 1808 [7].

MADAN (A. C.), *Living speach in central and south Africa.* Oxford, 1911.

MEINHOF (C.), *Grundriss einer Lautlehre der Bantusprachen.* 2e éd. Berlin, 1910 [8].

— *Grundzüge einer vergl. Grammatik der Bantuspr.* Berlin, 1906.

— *Sprachen des dunkeln Weltteils.*

MENDELSSOHN, *South African bibliography.* Londres, 1910.

SACLEUX (le P.), *Essai de phonétique*, Paris. 1905.

Le verbe « être » dans les langues bantoues, MSL. T. XV.

SCHLEICHER (A. W.), *Das persönliche Pronomen der Bantusprachen.* WZKM, VII.

— *Afrikanische Petrefakten.* Berlin, 1891.

STAPLETON (W. H.), *Compar. Handbook of Congo languages.* Yakusu, 1903.

STARR, *A bibliography of Congo languages*, 1908.

TORREND (le P.), *Comparative grammar of the South African Bantu languages*, Londres, 1892 [9].

— *Nouvelles études bantoues.* St. Glott. It., 1907.

VIVIEN DE SAINT-MARTIN, *Rapport annuel* Bull. Soc. Géographie, 1846.

WERNER, *Notes on Clicks in Bantu languages.* JRAS II.

— *The Bantu languages.* Londres, 1919.

WAUTERS ET BUYLE, *Bibliographie du Congo.* Bruxelles, 1895.

NOTES

1. Ouvrage qui a servi de base aux études postérieures ; l'auteur expose les faits hottentots à côté des faits bantou.

2. Exposé succinct des connaissances linguistiques de l'époque.

3. L'auteur reproduit un certain nombre de mots en un très grand nombre de dialectes, et essaye de déterminer ainsi un classement généalogique.

4. Très important ouvrage renfermant : 1° des vocabulaires de 276 parlers bantou et de 24 idiomes semi-bantou, disposés de manière à faciliter les études de phonétique comparée ; 2° quelques renseignements sur les affixes morphologiques de chaque idiome ; 3° une bibliographie très détaillée, bien qu'elle ne prétende pas être complète. L'auteur a essayé de classer tous les parlers étudiés.

5. Recueil de 246 vocabulaires recueillis auprès d'esclaves venus de l'intérieur, etc.

6. Dans la préface, l'auteur expose toutes les questions se rapportant aux langues africaines et l'état des connaissances contemporaines.

7. L'auteur fut peut-être le premier à signaler les ressemblances présentées par la plupart des langues parlées au Sud de l'équateur

8. Les travaux de M. Meinhof ont attiré l'attention de nombreuses personnalités allemandes et anglaises sur les langues bantou et la voie indiquée par lui a été suivie par une foules de missionnaires, de fonctionnaires, etc.

L'auteur de ce chapitre toutefois n'a eu connaissance de ces travaux qu'après qu'elle se fut engagée dans la voie suivie par elle depuis le début de ses études ; tout en rendant hommage à l'œuvre accomplie par M. Meinhof, elle tient à spécifier qu'il n'a jamais été pour elle ni un maître, ni un guide.

9. Ouvrage renfermant un grand nombre d'exemples empruntés á des idiomes variés et qui permettent de se rendre compte des particularités morphologiques des langues bantou.

ÉNUMÉRATION DES PRINCIPAUX PARLERS BANTOU ET BIBLIOGRAPHIES PARTICULIÈRES.

Le classement des langues bantou d'après leurs affinités linguistiques a été entrepris par Sir Harry Johnston qui énumère 366 parlers bantou et 87 parlers voisins qu'il appelle semi-bantou. Parmi ce nombre, certains ne sont connus que par quelques mots cités dans l'ouvrage de Koelle (voir ci-dessus) ; d'autres présentent bien quelques traits particuliers, mais ceux-ci correspondent souvent aux petits faits qui différencient les patois de villages d'une même région française et ne paraissent pas mériter le nom de dialectes distincts. Aussi a-t-il paru sage de ne donner en principe ici que les noms de langues connues par des ouvrages sérieux. Le classement géographique a été adopté ; car d'une façon générale il correspond à ce que donnerait un classement linguistique et il a l'avantage de ne pas induire en erreur en faisant croire à des parentés peut-être non existantes. (On trouvera sous chaque nom de groupe, une liste de noms de dialectes importants connus par des vocabulaires, etc.)

1. *Groupe nord ou Ganda*, au Nord-Ouest du lac Victoria Nyanza.

GANDA

Wilson, *Grammar of the Lu-ganda language*. Londres, 1882.

Missionary of the CMS, *Elements of lu-ganda grammar*. Londres, 1902.

Pilkington, *Lu-ganda-engl. and engl.-luganda vocabulary*. Londres, 1899.

NYORO

Maddox, *Elementary Nyaro grammar*. Londres, 1902; dialecte parlé au Sud-Est du lac Albert.

KEREWE

Hurel, *Essais de Grammaire* Kerewe. MSOS, 1909 (voy. Anthr., VI).

Dans ce groupe on rencontre un préfixe augmentatif et les noms de langues prennent le préfixe attestant *do (*ru*, *lu*, etc.); d'autre part le préfixe attestant *bo (*wu*, *bu*, etc) s'emploie pour le pluriel des diminutifs. Ces dialectes se distinguent donc principalement par leur système de préfixes. La correspondance $h = p$ des autres langues, est le trait le plus saillant de leur phonétique.

2. *Groupe Ruanda*, au Sud-Est du groupe ganda au Nord-Est du Tanganyika.

RUANDA

Hurel, *La langue Kinyarwanda*. MSOS, XIV.

KIRUNDI

P. van der Burgt, *Dictionnaire français-rundi*. Bois le Duc, 1903.

Éléments d'une grammaire kirundi. MSOS, V.

Menard, *Grammaire kirundi*, 1910.

Dictionnaire français-kirundi et kirundi-français. 1910.

Ce groupe présente certaines affinités avec le groupe ganda mais s'en sépare par l'emploi de *ki-* pour former les noms de langue.

3. *Groupe du Nord-Est* : région du massif du Kilima Njaro ; s'étend au Nord jusqu'aux limites du domaine bantou.

KIKOUYOU au Sud du Mont Kénia sur la rive gauche du Tana supérieur.

A. W. Mac Gregor, *English-kikouyou vocabulary*. 1904.

A. Hémery, *Engl.-kikouyou handbook*. 1903.

KAMBA entre le Tana moyen et le Kilima-ndjaro ;

T. Brutzer, *Kambasprache*. MSOS, 1906.

Mrs Hynde, *Kamba and Kikuyu vocabularies*. 1904.

TCHAGGA langue du Kilima-ndjaro, étroitement apparentée aux deux suivantes : mochi et siha.

Seidel, *Grammatische Elemente des Kichaga*. ZAOS, 1895.

MOCHI } (v. ci-dessus) langues du Kilima-ndjaro.
SIHA } Fokken, *Das Kisiha*. MSOS, 1905.

Ces dialectes présentent quelques traits phonétiques particuliers ; ainsi on y rencontre *ü* et *þ*.

4. *Groupe nord de la côte orientale et pays limitrophes.*

TAVETA-TAITA à l'Est du Kilima-ndjaro,

Hémery (Père), *Vocabulaire français-swahili-Taita.* Zanzibar, 1901.

Woodward (H. W.), *Ki-taita.* Zeitschr. f. Kolon. Spr., Berlin, 1914.

TCHASSOU dans les Monts Pare.

T. Kotz, *Grammatik des Tchassu.* Berlin, 1909.

POKOMO sur le Tana inférieur.

Würtz, *Grammatik des Pokomo.* ZAOS, 1895.

NIKA entre le Sabaki et l'Oumba, à l'Ouest de la région swahili ; les dialectes GIRYAMA et DIGO sont étroitement apparentés au Nika (mot qui désigne plutôt une région qu'un peuple).

Sparshott, *Nika-Engl. dictionary*, 1887.

Taylor (W. T.), *Giryama Vocabulary and Collections*, 1891.

Meinhof (C.), *Digo.* MSOS, 1905.

SHAMBALA dans l'Ousambara au nord du Pangani.

Rœhl (K.), *Grammatik der Schambalasprache*, 1911.

Werner (A.), *Notes on the Shambala Language.* JRAS., 1906.

BONDEI au Nord du Roufou inférieur.

Woodward, *Bondei Handbook.* Zanzibar. 1882.

ZIGOULA entre le Roufou inférieur et les Monts Ousagara.

Woodward (H. W.), *Handbook of the Zigula language.* Oxford, 1902.

SWAHILI. Idiome de Zanzibar et de la région côtière entre le Pagani et Lindi. Les Arabes et les Indiens avaient des comptoirs à Zanzibar lorsque Vasco de Gama y arriva et la langue swahili a été écrite avec les lettres arabes depuis quelques siècles ; on trouve donc une langue littéraire, voire même une langue poétique (le ki-ngozi), et des dialectes estimés vulgaires et propres à chaque localité. Les traitants arabes parlaient swahili, et ont répandu la langue jusqu'au Congo comme lingua franca. C'est un idiome fortement altéré au point de vue phonétique, car la sonore intervocalique attestant **d* a disparu. Ex :

Swah. *kua* =	bondei,	xosa,	ganda,	kimboundou	*kula*
makaa =	—	—	—		*makala*
amua =	—	—	—		*lamula*

Le vocabulaire renfermant de nombreux emprunts, dont quelques-uns anciens, se différencie passablement de celui des autres langues bantou.

Les travaux sur le swahili sont très nombreux tant en français qu'en anglais et en allemand. Nous renvoyons le lecteur aux bibliographies et n'en citerons qu'un en français et un en anglais.

P. Sacleux, *Grammaire des dialectes swahilis.* Paris, 1909.

Madan (A. C.), *Handbook of the swahili language.* Londres, 1894.

Voir en outre Werner (A.), *Catalogue of linguistic works.* JRAS., 1908.

COMORES. Les dialectes des îles Comores connus sont l'*angazija* et l'*anzuani* ; ils ont été assez peu étudiés jusqu'ici ; ils se rapprochent du swahili.

Heepe, *Die Komorendialekte Ngazidja u. Nzwani*. Hambourg, 1920.

ZARAMO au Sud de Dar-es-salam.

A. Worms, *Grammatik des Kizaramo*. ZAOS, 1897.

KAMI à l'Ouest d'Ou-zaramo.

Velten, *Ki-kami*. MSOS, 1900.

MATOUMBI au Sud de Zaramo près de Mohoro.

Krum, *Kimatumbi*. MSOS 1912, 1913.

5. *Groupe d'Afrique orientale.*

Ce groupe comprend les langues parlées entre le lac Tanganyika á l'Ouest, le lac Victoria Nyanza au Nord, le lac Nyassa au Sud et la région côtière. Les différences entre elles sont d'ordre phonétique ou syntaxique ; les morphèmes importants sont les mêmes partout, et le vocabulaire ne varie guère, quant au fond.

NYAMWZI au Sud du lac Victoria Nyanza; une bibliographie détaillée des dialectes de cette langue a été publiée par Struck (Mitt. a. d. Deutschen Schütz gebieten, 1910) ; les plus importants ou les mieux connus sont le nyanyembe, le soukouma, le kirwana, le kisumbwa, le kikonongo.

P. Capus, *Grammaire de shi-sumbwa*. ZAOS, IV.

— *Dictionnaire shi-sumbwa-français*. Paris, 1901.

NYATOUROU entre le Boubou et l'Ouambaere au Sud-Ouest d'Ou-soukouma.

Schregel (W.), *Grammatik der Kinyaturusprache*. MSOS, 1913.

GOGO entre le pays Nyatourou et la rivière Nyombe.

Nouveau Testament. Londres, 1899.

HEHE autour d'Iringa et dans le pays au Nord-Est du lac Nyassa.

Velten (C.), *Die Sprache der Wahehe*. MSOS, 1899.

POGORO entre Iringa et Mahengué au Sud du Rouaha.

Sachau (E.), *Pogoro*. Archiv f. d. Stud. deut. Kol. Spr., 1907.

KAGOUROU dans le pays d'Ou-sagara au Nord-Est de Pogoro et au Sud-Est de Gogo.

Last (J. T.), *Grammar of the Kaguru language*. Londres, 1886.

KONDE au Nord du Nyassa. MSOS, 1899.

Schumann, *Kondesprache*. MSOS, 1899.

YAO, langue d'un peuple jadis très puissant, établi à l'Est du lac Nyassa.

Hetherwick, *Handbook of the Yao language*. Londres, 1889.

NGANJA au Sud du Nyassa autour de Blantyre; on en connait plusieurs dialectes : *manganja*, *chi-nganja*, etc.

Riddel, *Grammar of Chi-nganja*. Édimbourg, 1880.

SENA } sur le Zambèze, deux dialectes très voisins et se rapprochant du
TETE } Nganja

V. d. Mohl (A.), *Grammatik der Bantu Sprache v. Tete*. MSOS, 1904.

Anderson (W. G.), *Grammar of the Sena language*. Londres, 1897.

CHWABO ou QUILIMANE.

Torrend, *Contes en Chwabo*. ZAOS, 1895, 1896.

6. *Groupe du Sud-Est* (Afrique orientale portugaise), Nord et Nord-Est de l'Union Sud-Africaine.

Ce groupe qui répond à une unité linguistique comprend deux sous-groupes : côtier et tchouana. Le premier comprend le makoua et le thonga (jonga, ronga, gwamba) de Lourenco-Marquès. Le second, le tlapi, le sotho le pédi, le kololo etc. Entre les deux se place le venda. Dans ces langues l'occlusive sourde *t* des autres langues est représentée, entre voyelles, par *r* et les consonnes occlusives sourdes sont généralement aspirées. D'autre part les occlusives palatales sont articulées d'une façon particulière et dans le groupe tchouana on trouve des mots à sons claquants ou inspirés vraisemblablement empruntés au hottentot, tandis que dans le thonga on signale une labiale palatalisée. Certains préfixes manquent dialectalement (ainsi *lo* et *le* sont représentés par *le* en sotho) et le pronom de la 2e pers. du pluriel a une forme particulière dans le groupe tchouana (*lo* ou *le*).

Ex. Makoua *raru*, « trois » = sotho *raro* = yao, ganda, zoul. etc. *tatu*
— *nrama*, « joue » = — *le rama* = — — — (li) *tama*

Dans le sous-groupe tchouana les groupes *mb*, *nd*, *ng* des autres langues sont inconnus et représentés par *p*, *t*, *k* (vraisemblablement sourdes douces) tandis que *mp*, *nt*, *nk* sont représentés par p^h, t^h, k^h (la nasale n'étant maintenue que lorsque le mot est monosyllabique Ex.

Sotho *bopa* = konde *bomba* « plâtrer ».
— *lakatsa* = zoulou *langaza* « convoiter ».
— *ota* = — *onda* « maigrir ».
— *motho* = — *muntu* « homme ».

Des phénomènes semblables ne se retrouvent que dans le groupe nord-ouest et dans les langues de la famille ouest-africaine et il est probable que les langues du groupe sud-est représentent des idiomes de tribus de même souche que ceux du Gabon et du Cameroun.

MAKOUA au Sud-Ouest du Mozambique.

Ch. Maples, *Handbook of the Makua language*. Londres, 1879.

RONGA ou THONGA région autour de Lourenço-Marques.

Junod (H.), *Grammaire ronga*. Lausanne, 1896.

GWAMBA ou HLENGWE au Sud des monts Zoutpans sur la rive droite du Limpopo inférieur.

Berthoud (P.), *Leçons de ši-gwamba*. Lausanne, 1883.

CHOPI ou LENGE, entre le Cap Corrientes et le Limpopo.

Smyth and Matthews, *Vocabulary... Grammar of Xilenge*. Londres, 1902.

VENDA entre les Monts Zoutpans et le Limpopo sur les limites du Transvaal et de la Rhodésia.

Meinhof (C.), *Das Tši- venda*. Leipzig, 1901.

SOTHO (ou *Se suto*) dans le Basouto-land dans le N. E. de la Colonie du Cap.

Casalis, *The Basuto language*. Paris, 1841.

Jacottet, *Grammar of the sesoutho langage*. Morija.

Kruger (H.), *Steps to learn the Sesuto language*. Morija, 1905.

Endemann (K.), *Wörterbuch der Sotho Sprache*. Hambourg, 1911.

KOLOLO, dialecte parlé par les descendants des Basotho qui ont fondé vers 1825 un royaume important sur le Zambèze à l'Ouest des chutes Livingstone. Ce dialecte se différencie nettement aujourd'hui du sotho et permet de déterminer les changements introduits au cours d'un siècle.

Dr Colyer, *Sikololo Grammar and Vocabulary*. Londres, 1917.

TCHOUANA dans le Betchouanaland.

Crisp (W.), *Notes towards a Secoana Grammar*. Londres, 1900.

Brown (J.), *Secwana Dictionary*. Londres, 1895.

7. *Groupe zoulou.*

Dans ce groupe on rencontre quelques mots à consonnes initiales inspirées; les préfixes ont presque toujours une voyelle initiale qui est la même que celle qui suit la consonne, ex. *abantu*, *umuntu*, *isisila*, etc. Les sourdes occlusives sont aspirées. Au cours du XVIIIe siècle des bandes de Zoulou ont traversé la Rhodésie et sont arrivées jusqu'au Nyassa ; leur route est jalonnée aujourd'hui par des colonies d'anciens Zoulou parlant des dialectes zoulou (tebele, ngoni) qui se différencient nettement de ceux des autres peuples de la région.

CAFRE ou XOSA au sud du Drakensberg dans la Cafrerie. Les travaux sur ce dialecte sont parmi les plus anciens que nous ayons sur des langues bantou ; l'habitude prise par les premiers écrivains d'écrire en un mot une forme verbale comprenant plusieurs particules donne aux textes cafres un aspect particulier tout superficiel.

Sparrman (A.), *Voyage au Cap de Bonne-Espérance*. 1776 (trad. franc., 1787).

Boyce, *Grammar of the Kafir language*, Graham'stown, 1834.

Mc Laren (J.), *Grammar of the Kafir language* Londres, 1906.

Kropf (A.), *Kafir-English dictionary*. Lovedale, 1899.

ZOULOU dans le Zoulouland. Le zoulou ne se différencie guère du cafre, mais le rôle politique joué par ce peuple guerrier et son extension ont donné à ce dialecte une situation prépondérante.

Schreuder, *Grammatik of Zulu Sproget*. 1850.

Bryant (A.), *Zulu-English dictionary*. Natal. 1905.

Grout (L.), *The isi-zulu* (grammaire). Londres, 1893.

TEBELE dans le Matabele land.

O'Neill (J.), *Shindebele Dictionary*. Londres, 1913.

NGONI au Nord-Ouest du lac Nyassa.

Elmsley, *Grammar of the Ngoni language*. Aberdeen, 1891.

8. *Groupe du Centre.*

Ce groupe géographique, qui présente d'ailleurs une certaine unité linguistique, comprend les parlers de la région située au Nord du Zambèze et à l'Ouest du Nyassa et Tanganyika. Les préfixes sont clairs. Les occlusives sont nettes ; les consonnes labiales seules sont faibles.

TABWA sur les rives sud-ouest du Tanganyika.

de Beerst, *Essai de grammaire Tabwa.* ZAOS, 1896.

van Acker, *Dict. kitabwa-français et français-kitabwa.* Bruxelles, 1907.

BEMBA au Sud du Tanganyika.

Schœfier, *Grammar of the Bemba language.* Oxford, 1907.

MAMBWE au Sud-Est du Tanganyika.

Jones, *Outlines of ki-mambwe grammar.* Londres, 1893.

BISA ou WISA à l'Est du lac Bangouéolo,

Madan (A. C.), *Introduction to the Wisa dialect.* Oxford, 1906.

LALA-LAMBA entre le Bangouéolo, les monts de Moutchinga et le Kafoué.

Madan (A. C.), *Lala-lamba Handbook*, Oxford, 1908.

Lala-lamba dictionary, Oxford, 1913.

SENGA à l'Ouest de Fort-Jameson.

Madan (A. C.), *Senga Handbook.* Oxford, 1905.

TONGA au Nord de Kazoungoula sur le Zambèze.

Torrend, *Compar. grammar of the South Afr. Bantu languages.* Londres, 1891.

Griffin, *Chitonga Vocabulary.* Oxford, 1915.

SOUBIYA à l'Ouest de Kazoungoula sur le Zambèze.

Jacottet (E.), *Langues du Haut-Zambèze.* Paris, 1896.

8. *Groupe occidental.*

Ce groupe comprend les langues parlées à l'Ouest du désert de Kalahari et du Zambèze. Un trait phonétique caractéristique du groupe est l'assimilation des voyelles et de certaines sonores à distance. Ex. :

Louyi *mana, mina* = *mala, mila* (langues nord, est, etc).

na tundu « il est sorti » infin. *ku tunda.*

na mono « il a vu » *ku mona.*

na singi « il a vu » *ku singa.*

HERERO dans le Hereroland au Sud du lac Etocha et au Nord du pays des Hottentots. Ce dialecte est remarquable par la présence de *ũ*, inconnu à la majorité des parlers bantou, et par la fréquence des faits d'assimilation de voyelles.

Viehe (G.), *Grammatik des Otji-herero.* Berlin, 1897.

NDONGA ou AMBO dans l'Ovamboland au Sud du Counéné (voir ci-dessous).

KOUANYAMA, au Nord du Counéné.

Brincker, *Lehrbuch des Oshi-kuanjama, ōshi-ndonga*, etc. Berlin, 1891.

NYANEKA, région de Huilla.

Lang, *Ensaios de Grammatica Nyaneka*. Lisbonne, 1906.

PP. du Saint-Esprit. *Diccionario Portuguez-Olulnyaneka*. Huila, 1896.

MBOUNDA au Nord-Est de l'Okavango.

Thomas (A. W.), *Comparative vocabulary of Sikololo-Silui-Simbunda*. Londres, 1916.

LOUYI dans le pays Barotse, autour de Lealoui.

Jacottet (E.), *Études sur les langues du Haut-Zambèze*. Paris, 1901.

UMBOUNDOU ou NANÓ, région de Benguella.

Stover (W.), *The structure of Umbundu* Boston, 1885.

Landers and Fay, *Vocabulary of Umbundu language*. Boston, 1885.

Pereiro do Nascimento, *Grammatica do Umbundu*. Lisbonne, 1894.

KIMBOUNDOU, région de Loanda (N.B. : le préfixe *ki* employé par Chatelain permet de distinguer cette langue de l'Umboundou).

De Cannecattim, *Diccionario da lingua Bunda*. Lisbonne, 1804.

— *Collecçao de observaçoes grammaticaes* Lisbonne, 1805.

Chatelain (E.), *Grammatica Elementar de Kimbundu*. Genève, 1889.

LOUNDA dans l'Angola.

de Carvalho, *Lingua da Lunda*. Lisbonne, 1889.

LOUBA ou LUA dans la province belge de Katanga.

On connait plusieurs dialectes de la langue louba, mais les différences ne sont pas grandes ; c'est peut-être la langue parlée avec le moins de variations dans une région aussi étendue.

Le Clercq (A.), *Grammaire de la langue des Bena Lulua*. Bruxelles, 1897.

Morrison (W. M.), *Buluba-Lulua Grammar and Dictionary*. New York, 1906.

Jenniges, *Dictionnaire Français-Kiluba*. Bruxelles, 1909.

9. *Groupe congolais.*

Ce groupe essentiellement géographique comprend les langues parlées dans les colonies belge et française dans les régions circonscrites par le Congo ou en avoisinant la rive gauche. Le dialecte de San Salvador fut le premier parler bantou décrit par un Européen.

Un sous-groupe linguistique congo est représenté par les noms de dialectes à radical *-ongo* mais avec des préfixes différents : *kišikongo*, *kikongo*, *kakongo* *luango* ou *yombe*. Dans ce groupe la dentale sonore est attestée par un son très voisin de *d* devant une voyelle fermée ; ainsi le préfixe *li* des langues du Nord, centre, etc., est représenté par *di*. On trouve un préfixe diminutif *fi*; la correspondance interdialectale $p = v$ est fréquente. L'ensemble des préfixes est moins régulier que dans les groupes précédents.

Sous-groupe congo.

Brusciotto, *Regulæ quædam pro Congensum idiomates captu*. Rome, 1659 (trad. angl. Guinness) Londres, 1882.

Bentley, *Dictionary and Grammar of the Kongo language*, 1887.

Vissey, *Grammaire fiote et Dictionnaire*. 1889 (dial. de l'estuaire).

— *Dictionnaire français-fiote*; *fiote-français*. 1889, 1890.

Butaye, *Grammaire Kikongo*. 1901.

Anonyme, *Dictionnaire français-congo*. 1772.

Cambier, *Essai sur la langue congolaise*. Bruxelles, 1890.

Laman, *Lärobok i Kongosproket*. Stockholm, 1912.

Starr, *A bibliography of Congo languages*, 1908.

Le Clercq, *Grammaire du Kiombe*. Anthropos, 1907, 1908.

Marichelle, *Dictionnaire Vili-Français*. Loango, 1902.

— — *Français-Vili*. Loango, 1912.

— *Etude du dialecte Vili*. Loango, 1913.

Kanioka parlé entre la Louloua et le Kassai.

Le Clercq, *Éléments de la langue Kanioka*. Paris, 1900.

— *Vocabulaire Français-Kanioka*. Paris, 1901.

Koutsou au Nord de la province de Sankourou (Congo Belge).

Van Hove, *La langue des Wankutšu* Anthropos, 1911.

Kele à l'Ouest de Stanleyville.

Stapleton, *Comparative Handbook of Congo languages*. Londres, 1903.

Lolo Nkoundou entre l'Équateur et la Boussira ; langue très répandue dont plusieurs dialectes ont été notés.

Anonyme, *A guide to the Lunkundu Language*. Londres, 1893.

10. *Groupe du Nord-Ouest.*

Ce groupe comprend les idiomes parlés au Nord du Congo sur la côte entre le Congo et le Wouri (Douala), et à l'intérieur au Nord, mais quelquefois aussi au Sud ; nous n'énumérons ici que les idiomes qui nous semblent être nettement bantou et donnerons plus loin la bibliographie de ceux qui nous paraissent appartenir à la famille ouest-africaine. Les parlers du groupe nord-ouest n'ont pas tous les préfixes nominaux des langues du Sud ni tous les suffixes verbaux, toutefois on en trouve des traces nettes ; le dialecte galoa présente l'alternance *l-r*, que l'on ne rencontre ailleurs que dans le groupe du sud-est. On peut rattacher à ce groupe les dialectes bangala et bangui, mais ces noms ont été employés pour des idiomes divers et les auteurs sont loin d'être d'accord ; n'ayant pu faire d'enquête sur place nous ne pouvons prétendre faire la lumière.

Bangui ; un dialecte bien connu de ce nom est parlé dans le pays appelé Bayanzi entre le lac Toumba et la jonction du Congo et du Kassaï.

Whitehead (J.), *Grammar and Dict. of the Bobangi language*, 1899.

TEKE, au Nord de Brazzaville.

Calloc'h, *Vocabulaire français-ifumu* (Bateke). Paris, 1911.

KALAI ou KÉLÉ, au Sud de l'Ogooué et à l'Est de Lambaréné.

Preston and Best, *Grammar of the Bakele language.*

GALOA, surtout au Nord de l'Ogooué et à l'Ouest de Lambaréné. Ce dialecte s'appelle quelquefois MPONGWE.

Le Berre, *Grammaire de la langue pongouée.* Paris, 1873.

Dictionnaire Mpongoué-français et franç.- Mpongoué, 1881.

DOUMA, au Nord du galoa, entre Lambaréné et l'Ouroungou.

Reeb, *Essai de Grammaire Douma.* Paris, 1895.

Dahin, *Vocabulaire Aduma-Français*, etc. Kempfen, 1895.

BENGA, autour de la baie de Corisco.

Mackey (J.), *Grammar of the Bengua language.* New York, 1855.

NOHO à l'embouchure de la Sanaga.

Adams, *Die Banoho.* MSOS, 1907.

BASA, au Sud-Est de Douala (Cameroun).

DOUALA à l'embouchure du Wouri (Cameroun).

Saker, *Elements of the Dualla language.* 1855.

Dinkelacker, *Wörterbuch der Duala Sprache*, Hambourg, 1914.

BOUBÉ, ensemble des dialectes parlés dans l'île de Fernando Po.

Clarke, specimens of dialects... of West Africa Londres, 1841.

— *Introduction to the Fernandian Tongue.* Berwick, 1848.

Juanola, *Primer Paso a la Lengua Bubi.* Madrid, 1890.

KWIRI, sur les pentes du Mont Cameroun.

Schuler (E.), *Die Sprache der Bakwiri.* MSOS, XI.

ISOUBOU, au Nord-Ouest du Mont Cameroun.

Merrick, *Dictionary of Isubu.* 1842.

LA FAMILLE OUEST-AFRICAINE

La famille linguistique ouest-africaine, étroitement apparentée à la famille bantou, comprend un grand nombre de langues parlées au Nord du Congo, au Gabon, au Cameroun, au Soudan et tout le long de la côte sud de l'Afrique occidentale.

Il est difficile de dire si les traits particuliers de ce groupe sont dus à des influences étrangères exercées sur des langues bantou, ou s'ils représentent une évolution propre d'un dialecte africain voisin du bantou commun.

Les traits particuliers qui différencient cette famille de sa voisine bantou sont :

1° la présence de syllabes fermées et de nombreux mots monosyllabiques à consonne finale. Ex. : boulou *mot* « homme », *nkom* « prisonnier », *nkes* « oiseau ».

2° des préfixes moins nombreux, et un système de préfixation moins régulier.

3° des formes verbales très distinctes, bien qu'obtenues également par suffixation et par adjonction de particules préposées. Ex. : boulou *ma boya esae* « j'ai fait l'ouvrage » ; *me boban esae* « j'ai fait l'ouvrage il y a plusieurs jours » ; *bo* « faire ».

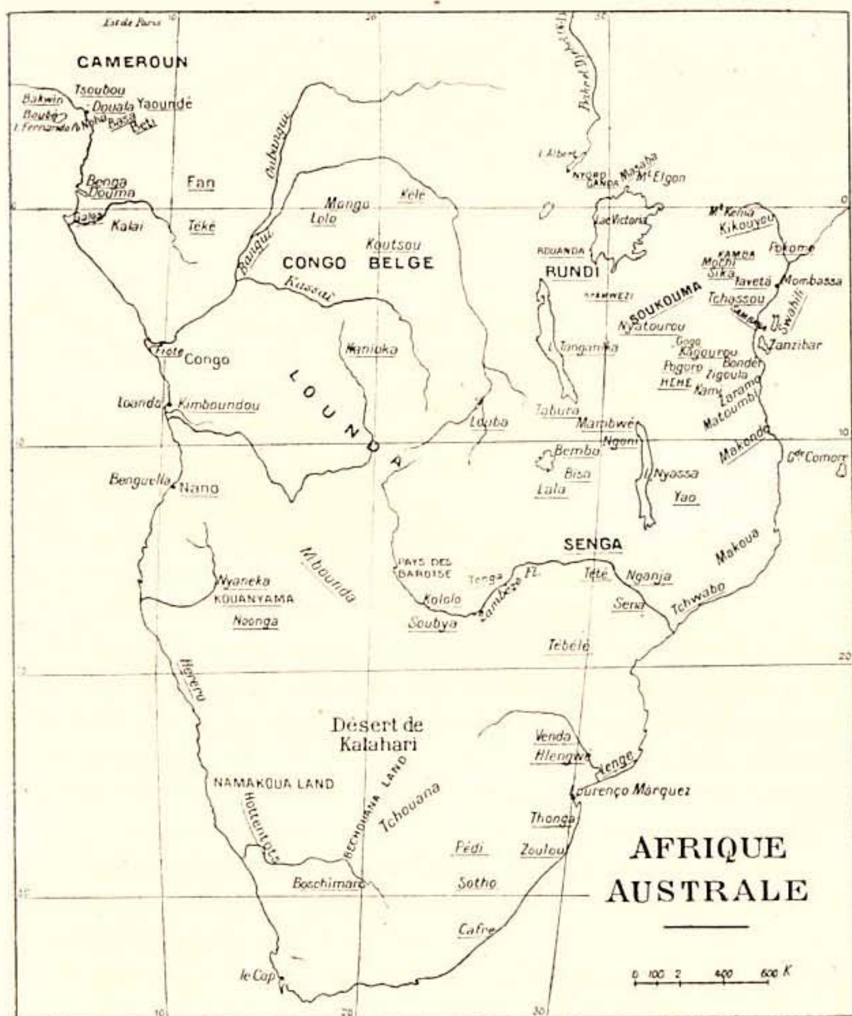
Beaucoup des parlers de cette famille ayant été compris par M. Delafosse dans le chapitre précédent nous n'indiquerons ci-dessous que les noms des dialectes appartenant au groupe fan.

BIBLIOGRAPHIE

V. Largeau, *Encyclopédie pahouine*. Paris, 1901.
H. Nekes, *Lehrbuch d. Jaunde-Sprache*. Berlin, 1911.
Tronje v. Hagen, *Lehrbuch d. Bulu Sprache*. Berlin, 1914.

Note. Ces ouvrages permettront au lecteur curieux de se rendre compte des traits propres du groupe. Il n'a pas paru utile de signaler ici tout ce qui a été écrit sur le pahouin, ni de faire entrer dans cette liste des descriptions des langues du Cameroun dont la place exacte dans les groupements africains est encore douteuse.

L. Homburger.



CORRECTIONS. — En haut à gauche, lire *Isoubou* et non *Tsoubou*
Au Sud du lac Tanganika, lire *Tabwa* et non *Tabura*.

LES LANGUES BOCHIMANES ET HOTTENTOTES

Les langues bochimanes sont peu connues, car les négrilles ou pygmées qui les parlent et qui habitent la forêt africaine et les régions désertiques du sud de l'Afrique emploient souvent les parlers des populations sédentaires voisines.

Quelques voyageurs pensant décrire un parler de négrilles n'ont fait que montrer les modifications apportées par eux au langage de leurs voisins ou maîtres, et ce n'est guère qu'au cours du XX[e] siècle qu'on commence à avoir quelques renseignements précis sur les parlers bochimans et à pouvoir affirmer avec certitude leur parenté avec les dialectes hottentots.

Les Bochimans (« hommes des bois » en hollandais) sont tout au plus quelque 50.000.

Les dialectes hottentots sont connus depuis longtemps. Les Hottentots peuplaient autrefois tout le pays situé au Sud du Cunéné et du Zambèze ; mais, refoulés par les Bantou venus du Nord et par les Européens de la Colonie du Cap, ils ne sont plus représentés que par quelques clans métissés, épars dans cette colonie, et par les Nama qui habitent aujourd'hui entre le 24° S. et le cours inférieur de l'Orange (environ 250.000 au total).

La langue des Nama a été décrite par plusieurs auteurs et généralement sous le nom de hottentot. Ce mot dérive du hollandais et désigne dans cette langue les personnes qui bégayent, qui articulent mal avec leur langue. Les Hottentots emploient en effet non seulement des sons articulés avec expulsion d'air, mais des sons accompagnés d'inspiration buccale ou succion, qui ont fait dire à Dapper que leur langage ressemblait plutôt aux cris des dindons qu'à des paroles humaines.

Le plus ancien texte habituellement cité est un vocabulaire donné par Pierre Kolbe en 1719, mais l'auteur de ce chapitre a

trouvé un petit vocabulaire hottentot dans le dictionnaire malgache de Flacourt publié à Paris en 1658 et un autre dans un ouvrage de Sir Th. Herbert (Londres 1638). La première description grammaticale parut en 1854 et depuis cette date l'idiome nama a été l'objet de travaux divers. La grammaire de Hahn est l'œuvre d'un missionnaire né et élevé parmi les Hottentots. Il rappelle dans son introduction qu'Hérodote et Pomponius Mela ont employé pour décrire la langue des troglodytes d'Éthiopie des expressions en tout point semblables à celle de Dapper et se demande si les Hottentots habitaient autrefois le Nord du continent. Il est vraisemblable que les Bochimans et les Hottentots représentent un seul groupe linguistique et que les tribus décrites par Hérodote étaient des négrilles à parler semblable.

Aujourd'hui les San ou Bochiman du Sud emploient sept articulations inspirées, tandis que les Hottentots n'en connaissent que quatre : une dentale, une palatale, une cérébrale et une latérale ; ces articulations inspirées ou clicks sont de véritables consonnes au même titre que les articulations expirées, mais ne se rencontrent qu'à l'initiale des mots [1].

Les consonnes expirées comprennent les occlusives *p*, *b*, *t*, *d*, *g'*, *k*, *g*, les nasales *m*, *n*, *ṅ*, les mi-occlusives, etc., *w*, *ts*, *s*, *z*, *ḥ*, *r*, *h*.

Les voyelles notées *a*, *e*, *i*, *o*, *u* peuvent être longues ou brèves, fermées ou ouvertes.

Les diphtongues et les voyelles nasalisées se rencontrent en nama.

Les mots homonymes se distinguent par le ton qui présente trois degrés : haut, normal, bas.

Le mot simple est monosyllabique ; les mots dissyllabiques ou polysyllabiques sont ou des composés ou des mots avec morphèmes suffixés.

Certains suffixes ne sont plus employés pour former des mots aujourd'hui, tandis que d'autres sont encore vivants.

1. Les seules qui figurent dans les exemples qui suivent sont notées ! et ǂ.

Le système des suffixes pronominaux différencie :

trois genres : masculin, féminin, commun ;

trois nombres : singulier, duel, pluriel ;

trois cas : dénominatif, intensif ou relatif, apostrophe.

A la première personne du duel et du pluriel il existe des formes distinctes pour l'inclusif et l'exclusif. Ex. :

celui-ci, *nēb* celle-ci, *nēs* ceci (com.), *nēi*.
celui-là, *nāb* celle-là, *nās* cela », *nāi*.

sa pipe, *eĩb khob* (d'une femme) *eĩs khob* (de quelqu'un), *eĩikhob*.
Le possessif peut être rendu par postposition de la particule *ã* [1] suivie du suffixe pronominal représentant le possesseur. Ex :

gomas ãb « sa vache » (d'un homme),
gomas ãs » » (d'une femme).

La suffixation des suffixes pronominaux à une racine peut donner des substantifs. Ex. :

!gãi « bon, » *!gãib* « le bon », *!oa* « contre », *!oab* « adversaire ».

Les formes verbales s'obtiennent par affixation de morphèmes variés qui, seuls ou en se combinant, permettent de différencier :
1° une conjugaison positive et une conjugaison négative. Ex. :

mũ « voir », *mũ-tama* « ne pas voir ».

2°	une forme simple				
		ou énumérative	*mũ-ta*.		
	»	»	stative		
		ou habituelle	*mũ ta hã*.		
	»	»	progressive	*mũ ta ra*.	
3°	»	»	primitive	*mũ*	« voir ».
	»	»	déterminée	*mũ ba*	« examiner ».
	»	»	causative	*mũ gei*	« faire voir ».
	»	»	réciproque	*mũ gu*	« se regarder l'un l'autre ».
	»	»	diminutive	*mũ ro*	« voir un peu».
	»	»	optative	*mũgao*	« visiter ».
	»	»	potentielle	*mũkha*	« pouvoir voir ».

1. La particule *ã* rappelle la particule *a* qui figure dans les pronoms possessifs africains et provient peut-être d'un emprunt.

4° un présent	*tita gye mũ*	« je vois ».
un prétérit	*tita go mũ*	« j'ai vu ».
un futur	*tita nĩ mũ*	« je verrai ».
5° un mode subjonctif	*tita gye ga mũ*	« que je voie ».
» » impératif	*nau re*	« écoute donc ».
» » infinitif	*mũ*	« voir ».
» » participe	*mũ !a*	« voyant ».
6° une voix active	*mũ*	« voir ».
» » passive	*mũ hê*	« être vu ».

Les qualificatifs sont soit des termes qualificatifs purs, soit des mots rendus qualificatifs par suffixation de certaines syllabes caractéristiques. Ex. :

gei	« vieillir »	*geira*	« vieux ».
aob	« homme »	*aore*	« mâle, masculin ».
gams	« eau »	*gamio*	« sans eau » [1].

En général, l'adjectif précède le substantif; il est postposé dans une phrase emphatique et prend alors le suffixe du substantif ou du pronom. Ex. :

tita geita « moi le grand », *aob geib* « homme grand ».

La numération hottentote est décimale.

Les rapports exprimés par nos prépositions le sont en hottentot par des particules postposées.

Dans son étude comparée des langues hottentotes et des dialectes bochimans voisins, Planert a mis en lumière des similitudes tellement nombreuses que le doute sur l'unité primitive ne peut subsister ; toutefois les renseignements sur les langues des autres négrilles étant peu nombreux il n'est pas possible de décrire ces idiomes dans leur ensemble et il ne semble pas utile de reproduire ici les détails donnés par Planert.

1. Ces suffixes rappellent les suffixes africains, en particulier *o* inversif en bantou, voir p. 568.

CLASSEMENT ET BIBLIOGRAPHIE DES LANGUES BOCHIMANES ET HOTTENTOTES

Les langues bochimanes et hottentotes, qui paraissent apparentées entre elles peuvent être divisées en trois groupes :

hottentot
bochiman
négrille.

Le groupe hottentot comprend plusieurs dialectes parlés dans la Colonie du Cap au Nord-Ouest et Ouest et un peu au delà des limites de la Colonie. Le dialecte qui a été le plus étudié est celui des Nama. Les principaux ouvrages sont :

Hahn, *Die Nama Sprache.* 1870.
J. Kroenlein, *Wortschatz der Khoi-Khoin*, Berlin, 1889.
Meinhof, *Die Nama Sprache.* Berlin, 1909.
Planert, *Über die Sprache der Hottentotten und Buschmänner.* MSOS, 1905.
Tindall, *Gram. and vocab. of the Namaqua.* Cape Town, 1857.

Les langues des Bochimans du désert de Kalahari et des régions limitrophes sont connues par des documents épars dans les récits de voyage et par des monographies dont voici quelques titres :

Bertin, *The Bushmen and their language.* Londres, 1885.
Vedder (H.), *Grundriss einer Grammatik der Buschman Sprache* Z. f. Kolonialsprachen I.
Wuras, *Outline of the Bushmen language.* Z. f. eingeborenen Sprache. 1920.

Les langues des négrilles de la forêt équatoriale ont été sommairement étudiées en même temps que leurs mœurs et leurs caractères anthropométriques par plusieurs voyageurs, mais il n'y a guère d'ouvrage linguistique important. On peut citer parmi les mémoires :

Ouzilleau, *La langue des Pygmées de la Sanga.* Rev. d'Ethnographie et de Sociologie, 1911.
Engels, *La langue des Ba-tua.* Revue Congolaise, 1912.

L. HOMBURGER.

LANGUES AMÉRICAINES [1].

Il est impossible de donner une description générale s'appliquant à toutes les langues américaines. Dans la belle préface qu'il a écrite pour le *Handbook of american indian languages* (**1**), Boas a montré à l'évidence que les caractères que l'on considérait autrefois comme spéciaux aux langues du Nouveau-Monde, alors qu'on n'en connaissait que quelques-unes, ont fait faillite, lorsque le champ de nos connaissances s'est élargi. Je ne puis mieux faire que de reproduire ici ce qu'a écrit à ce sujet le savant linguiste américain. Boas rappelle tout d'abord que, d'après la vieille classification, les langues américaines étaient caractérisées par le polysynthétisme et l'incorporation, puis il ajoute : « Voici, « par exemple, un mot eskimo qui indique bien ce que l'on entend « par polysynthétisme : *takusariartorumagaluarnerpâ*? pensez-vous « que réellement il ait l'intention d'aller s'occuper de cela ? (*taku-* « *sar*[*pâ*], il s'occupe de cela ; *-iartor*[*poq*], il va à ;*-uma*[*voq*], il a « l'intention de ;-[*g*]*aluar*[*poq*], il fait ainsi-mais ;*-ner*[*poq*], « pensez-vous il ; *-â*, interrogatif, 3me personne). On voit de « suite qu'il n'y a aucun rapport entre les éléments ajoutés au « thème fondamental et les éléments grammaticaux des langues « indo-européennes. Un exemple analogue nous est fourni par « le Tsimšian *t-yuk-ligi-lo-d'ᴇp-dāʟᴇt*, il commença à le déposer « quelque part à l'intérieur (*t*, il ; *yuk*, commencer ; *ligi*, quelque « part ; *lo*, dans ; *d'ᴇp*, en bas ; *dāʟ*, déposer ; *-t*, cela) ».

« Par incorporation, on entend que les langues américaines

1. Dans ce travail, le nom des langues ou dialectes éteints ou supposés tels est précédé d'un astérisque.

Les chiffres gras entre parenthèses renvoient à l'index bibliographique qui termine chacune des divisions de ce chapitre (Introduction, p. 605 ; Amérique du Nord, p. 605 ; Amérique centrale, p. 638 ; Amérique du Sud et Antilles, p. 701 ; L'écriture en Amérique, p. 712).

« ont tendance à incorporer le régime de la phrase, substantif « ou pronom, dans l'expression verbale. En voici des exemples : « Nahuatl : *ni-petla-tšiwa*, je fais des nattes (*petla-tl*, natte) ; « Pawni : *tA-t-ī'tka'wit*, je pioche de la boue (*tA*-, indicatif ; « *t*-, je ; *ī'tkār*", boue ; *-pīt*, piocher [la rencontre de *r* et *p* « donne '*w*]) ; Oneida : *g-nagla'-sl-i-zak-s*, je cherche un village « (*g*-, je ; *-nagla'*, vivre ; *-sl-* substantif abstrait ; *-i-*, forme « verbale ; *-zak*, chercher ; *-s* indique la continuité) ».

« Lorsqu'on étudie plus à fond la structure de beaucoup de « langues américaines, on s'aperçoit qu'il est faux de dire en « général de toutes ces langues qu'elles sont polysynthétiques et « incorporantes. Il y a en Amérique un nombre considérable de « langues, où les pronoms ne sont pas incorporés mais simple- « ment juxtaposés au verbe, et également beaucoup de langues, « où l'incorporation de nombreux éléments dans un seul mot ne « se produit presque pas. Le Tšinuk, par exemple, ignore la « polysynthèse. Très rares sont les cas, s'il y en a, où un seul « mot tšinuk exprime un réseau compliqué d'idées... L'étude « de la syntaxe du Tšinuk montre que les thèmes verbaux ne « sont modifiés que par les pronoms et quelques adverbes, et que « les substantifs n'ont presque aucune tendance à incorporer de « nouvelles idées, telles que celles exprimées par nos adjectifs. « D'un autre côté, l'Athapaskan, le Haida et le Tlingit peuvent « être cités comme types de langues, qui, bien que polysynthé- « tiques, n'incorporent pas réellement le régime, mais traitent le « pronom sujet et le pronom régime comme des éléments indé- « pendants ».

« Parmi les langues de l'Amérique du Nord, l'Iroquois seul « tend tellement à incorporer le substantif régime dans le verbe « et en même temps à modifier la forme indépendante du verbe, « qu'il peut être considéré comme une des langues typiques qui « incorporent le régime. Ce caractère se retrouve aussi à un « degré moindre dans le Tsimšian, le Kutenai et le Šošon ; il est « fortement marqué dans les langues Kaddo. L'Eskimo, l'Algon- « kin, le Kwakiutl se limitent à une incorporation plus ou « moins stricte du pronom-régime. En Šošon, l'incorporation

« du pronom et du substantif régimes est si peu prononcée qu'il « est à peu près indifférent de ranger ces formes parmi les formes « incorporées ou parmi celles qui ne le sont pas. Si nous consi- « dérons d'autres régions de l'Amérique, les mêmes phénomènes « apparaissent clairement, et il n'est pas possible de dire que ces « deux caractères soient typiques et se trouvent dans toutes les « langues américaines ».

Boas étudie encore quelques autres caractères et conclut qu'il y a des faits américains, c'est-à-dire se rencontrant fréquemment dans de nombreuses langues américaines, mais qu'il n'y a pas de caractères morphologiques communs à toutes ces langues.

Abordant ensuite la question au point de vue phonétique, il arrive à la même conclusion décevante.

Si donc, comme il est probable, il a existé un Américain commun, les langues qui en sont issues ont si fortement divergé qu'elles ne conservent pas de traces appréciables, dans l'état actuel de nos connaissances, de leur origine commune, et, comme l'histoire de ces langues nous échappe entièrement, les documents les plus anciens ne remontant pas, dans les cas les plus favorables, au delà du XVI[e] siècle, le problème paraît presque insoluble. Où en seraient les études indo-européennes si elles n'avaient eu d'autres bases que des vocabulaires, souvent très limités, et quelques grammaires ne remontant pas à plus de quatre cents ans ?

De plus, le fait que l'Amérique a été une terre si hospitalière aux colons blancs ou de couleur (actuellement, les 15.000.000 d'Indiens encore existants sont noyés dans une population totale de 197.810.070 âmes [1]) a eu pour conséquence la disparition rapide, soit par extinction, soit par absorption, d'un grand nombre de tribus indigènes ; beaucoup de langues ont ainsi cessé d'être parlées depuis la découverte, sans qu'on ait pu recueillir sur elles les documents nécessaires à leur étude, et trop souvent le linguiste doit se contenter de maigres listes de mots, dont la notation même est suspecte.

1. Non comprises la population du Honduras britannique pour laquelle je n'ai pas trouvé de chiffres se rapportant au pourcentage des Indiens, et la population des Antilles et de l'Uruguay, où l'élément indigène a pratiquement disparu.

Il semble en outre qu'en Amérique, plus que dans tout autre continent, les conditions aient été favorables à une différenciation profonde des langues. La densité de la population indigène y a toujours été très faible. Réparties sur de vastes territoires, les peuplades y ont vécu dans un isolement relatif qui a dû singulièrement hâter l'évolution divergente des dialectes de l'Américain commun.

Encore que les statistiques relatives au nombre des Indiens soient très souvent sujettes à caution (par exemple, le chiffre de la population indienne mexicaine varie, suivant les recensements, de 1.960.306 à 4.500.000), elles fournissent cependant d'intéressantes indications (**3**, II, 286-287, article *Population* ; **4** ; **5**).

On estime qu'au moment de la découverte, la partie de l'Amérique du Nord, qui s'étend au Nord du Mexique, était peuplée de 1.150.000 Indiens environ et que la densité de la population, suivant les diverses régions, était la suivante :

	NOMBRE D'INDIENS.	SUPERFICIE EN KM².	NOMBRE D'INDIENS PAR KM².
Alaska	72.000	1.530.327	0.05
Groenland	10.000	88.100	0.11
Possessions anglaises	220.000	9.788.120	0.02
États-Unis	846.000	7.839.064	0.11
Totaux	1.148.000	19.245.611	0.06

Actuellement, cette population est réduite à 403.000 âmes, dont 340.830 aux États-Unis et 14.354 au Groenland.

Le tableau suivant donne les renseignements que j'ai pu réunir sur l'Amérique centrale (à l'exception du Honduras britannique) et l'Amérique du Sud (Uruguay exclus).

Amérique centrale.

	NOMBRE D'INDIENS.	SUPERFICIE EN KM².	NOMBRE D'INDIENS PAR KM².
Mexique	4.500.000	1.985.200	2.27
Honduras	583.070 (Métis et Indiens)	114.670	5.08
Costa Rica	37.765	48.550	0.78

Panama	48.000	86.250	0.56
Nicaragua	211.062	128.340	1.64
Guatémala	350.000	113.030	3.10
Salvador	708.247	21.160	33.47
Totaux	6.438.144	2.497.200	2.6

Amérique du Sud.

	NOMBRE D'INDIENS.	SUPERFICIE EN KM².	NOMBRE D'INDIENS PAR KM².
Colombie	160.436	1.206.200	0.13
Équateur	500.000	307.243	1.63
Pérou	3.459.600	1.833.916	1.89
Bolivie	922.000	1.470.196	0.63
Chili	50.000	750.572	0.07
Argentine	38.000	2.950.520	0.01
Paraguay	130.000	253.100	0.51
Brésil	3.072.622	8.497.540	0.36
Vénézuéla	325.000	1.020.400	0.32
Guyane française	1.400	78.000	0.02
Guyane anglaise	8.000	234.380	0.03
Guyane hollandaise	2.700	129.100	0.02
Totaux	8.669.758	18.731.167	0.5

La population indienne actuelle de l'Amérique s'élève donc à environ 15.000.000 d'individus (exactement 15.510.902 ou 12.971.208, suivant que l'on adopte le chiffre de 4.500.000 Indiens pour le Mexique ou celui de 1.960.306 du recensement de 1910), répartis sur une superficie de 40.473.978 km² [1] ; la densité de la population indigène est donc de 0,4 par km² en moyenne.

Si l'on admet que la diminution du nombre des Indiens dans toute l'Amérique depuis la conquête est de l'ordre de la diminution calculée pour l'Amérique du Nord, on arrive à cette conclusion, qui n'a évidemment qu'une valeur toute relative, qu'au moment de la découverte, le Nouveau Monde devait avoir *au maximum* de 40 à 45 millions d'habitants, chiffre qui correspond

1. Ce chiffre représente la superficie totale de l'Amérique, diminuée de la superficie du Honduras britannique, des Antilles et de l'Uruguay.

à une densité moyenne de population extrêmement faible : 1 habitant environ par km².

Une dernière cause me semble expliquer, dans une certaine mesure, la multiplicité des langues américaines.

Alors que, dans le vieux monde, il y a eu des langues de civilisation, qui, en se substituant aux diverses langues indigènes, ont refait l'unité linguistique sur d'immenses territoires, en Amérique, ces langues de civilisation sont exceptionnelles, et quand elles ont existé (le Kitšua et le Nahuatl, par exemple, pour les temps précolombiens, le Tupi-Guaraní, à une époque récente), leur action n'a été ni assez durable, ni assez puissante pour anéantir le substratum linguistique des régions où elles ont pénétré.

C'est pour ces raisons multiples que l'on rencontre, en Amérique, un grand nombre de langues (123 d'après ma classification), entre lesquelles il a été jusqu'ici impossible de découvrir aucun lien de parenté, et dans lesquelles on ne peut même pas mettre en évidence des caractères généraux communs, ainsi que Boas l'a montré.

Est-ce à dire qu'on doive renoncer à établir quelque jour une classification linguistique américaine, plus simple que celle que je vais exposer ? Certes non. Les résultats remarquables obtenus par les linguistes en Amérique du Nord, notamment en Californie, autorisent l'espoir que, malgré des conditions si défavorables, on pourra peu à peu constituer des groupes plus importants par la réunion de langues qui, actuellement, paraissent étrangères l'une à l'autre; mais il est probable, sinon certain, qu'on ne pourra jamais pousser cette simplification aussi loin qu'on l'a fait dans les domaines indo-européen et africain.

Dans l'exposé qui va suivre, on s'étonnera peut-être que chaque famille linguistique ne soit pas caractérisée par les traits essentiels de sa morphologie grammaticale, suivant le plan adopté dans l'ensemble de cet ouvrage. La raison en est que les familles linguistiques américaines ont été constituées uniquement ou presque uniquement sur des similitudes de vocabulaires. Certes, il existe d'excellentes grammaires pour un assez grand nombre de

dialectes, mais il n'y a pas encore une vraie grammaire comparée d'une seule famille linguistique. C'est ainsi que les collaborateurs du *Handbook of american indian languages* (**1**) se sont contentés de prendre comme type de chaque famille un de ses dialectes les mieux connus. Si utile que soit la tâche qu'ils ont accomplie, elle ne doit pas cependant nous faire illusion. L'excellente grammaire du Hupa par Goddard, par exemple, est aussi peu une grammaire comparée de la famille Athapaskan qu'une bonne grammaire espagnole est une grammaire comparée des langues romanes.

Cette grave lacune de la linguistique américaine tient surtout à la valeur très inégale des matériaux que l'on possède sur chaque dialecte; elle provient aussi de ce fait que les langues américaines n'ont été l'objet de travaux sérieux que depuis peu d'années, et que, trop souvent, les personnes qui s'y consacrent n'ont ni la culture générale ni la discipline scientifique des linguistes qui se sont spécialisés dans l'étude des langues du vieux continent.

Le travail que je présente ici n'est donc qu'un inventaire, aussi complet que possible, des langues américaines. On n'y trouvera — et il ne pouvait pas en être autrement — aucune description de ces langues.

Pour chaque famille, j'ai donné tantôt une, tantôt plusieurs indications bibliographiques qui donneront la clef de la bibliographie complète de la famille. Ces références ne se rapportent donc pas nécessairement à l'ouvrage essentiel à consulter, mais au dernier travail paru, par lequel il est possible d'atteindre, avec un minimum de recherches, tous les documents connus à ce jour pour chaque famille.

Cet inventaire des langues américaines a déjà été fait plusieurs fois. Je ne citerai pas ici les ouvrages généraux sur les langues du monde, où les idiomes américains sont catalogués; je ne ferai mention que des trois ouvrages de Ludewig (**6**), de Brinton (**2**) et de Wissler (**8**), spécialement consacrés à l'Amérique dans son ensemble, que l'on peut avoir intérêt à consulter.

En raison de l'enchevêtrement sur le terrain des différentes familles, dont certaines se répartissent presque du Nord au Sud

d'un continent, il ne m'a pas paru possible d'adopter un ordre géographique, en dehors de la grande division de l'Amérique en trois parties : Amérique du Nord, Amérique centrale, Amérique du Sud et Antilles. Les familles sont donc rangées par ordre alphabétique.

Les noms que j'ai adoptés, pour ces familles et pour les groupes qui les composent, sont en principe transcrits dans l'écriture phonétique qui a été adoptée pour l'ensemble du présent ouvrage. Toutefois, certaines concessions ont été faites aux usages orthographiques des langues européennes (emploi de *th*, *kh*, des accents français et espagnols, etc.) ; quelques noms très usuels, passés dans les langues européennes, ont été conservés tels quels (Siou, Iroquois) ; des traductions européennes de noms indigènes (Blackfoot, Gros-ventres) ont été aussi occasionnellement adoptées. Sauf dans les noms nettement européanisés, la désinence du pluriel (*-s*, *-es*) a été systématiquement supprimée. Très souvent, les noms des parlers ne sont autres que ceux des tribus qui en font usage.

Pour l'Amérique du Nord et pour l'Amérique centrale, j'ai restreint beaucoup la synonymie, me contentant souvent de donner le nom adopté dans le *Handbook of american Indians* (3) ou dans le travail de Thomas et Swanton (7). Quand on ne trouvera pas à l'index du présent ouvrage un nom de langue ou de tribu de l'Amérique du Nord ou de l'Amérique centrale, on voudra bien se reporter à l'index qui se trouve à la fin de ces deux livres fondamentaux. On y trouvera réunis avec un soin méticuleux tous les synonymes connus.

Pour l'Amérique du Sud, comme il n'existe pas de livre de référence semblable, j'ai cherché à noter tous les synonymes usuels, mais j'ai dû renoncer à donner la liste de tous ceux qui ont été employés plus ou moins occasionnellement pour ne pas allonger démesurément l'index. Malgré cette lacune, le lecteur pourra cependant, je crois, retrouver assez facilement dans l'index, sous l'une des transcriptions que j'ai notées, les noms de parlers ou de tribus qu'il aurait relevés dans la littérature sous une autre forme, s'il veut bien se guider non sur l'orthographe mais sur la prononciation de ces noms.

INDEX BIBLIOGRAPHIQUE.

1. BOAS (Franz). *Handbook of american indian languages*. Bureau of american ethnology, Bulletin 40. Washington, t. I. 1911, t. II, 1922.

2. BRINTON (Daniel G.). *The american race : a linguistic classification and ethnographic description of the native tribes of North and South America*. New York, 1891.

3. *Handbook of american Indians North of Mexico* (edited by Frederick Webb HODGE). Bureau of american ethnology, Bulletin 30, 2 vol. Washington, 1907-1910.

4. JURASCHEK (I. von) et SCHULLERN (Hermann R. von). *Otto Hübners geographisch-statistische Tabellen aller Länder der Erde*. 66e édition, 1920. Vienne, 1921.

5. LINDQUIST (G. E. E.). *The red man in the United States ; an intimate study of the social, economic and religious life of the american Indian*, made under the direction of — (with a foreword by Honorable Charles H. BURKE, Commissioner of indian affairs). New York, 1923.

6. LUDEWIG (Hermann E.). *The literature of american aboriginal languages* (with additions and corrections by Professor Wm. W. TURNER). Trübner's Bibliotheca glottica, I. Londres, 1858.

7. THOMAS (Cyrus) et SWANTON (John R.). *Indian languages of Mexico and Central America and their geographical distribution*. Bureau of american ethnology, Bulletin 44. Washington, 1911.

8. WISSLER (Clark). *The american Indian, an introduction to the anthropology of the New World*. 2e édition. New York, 1922.

Nota.

Les cartes jointes à ce mémoire (planches 15, 16, 17 et 18) n'ont pas été exécutées avec le soin que j'eusse désiré. En les présentant au public, je tiens à décliner toute responsabilité personnelle en ce qui concerne la manière dont mon projet a été réalisé par le cartographe qui m'a été imposé.

I. LANGUES DE L'AMÉRIQUE DU NORD.

Pour établir la liste des familles linguistiques nord-américaines, je me suis servi presque exclusivement du travail de révision, fait il y a dix années, par P. E. Goddard (4). Je me suis contenté de le compléter sur certains points avec des renseignements tirés du *Handbook of american Indians* (5) et du *Handbook of american indian languages* (1). En outre, je l'ai mis à jour, en tenant compte des études parues depuis sa publication, qui ont modifié la classification admise par Goddard. Le lecteur, en se reportant à ces articles et à l'excellente bibliographie du mémoire de Goddard, aura la clef de toute la bibliographie linguistique nord-américaine.

Malgré les remarquables efforts des savants américains, le nombre des familles linguistiques nord-américaines s'élève encore au chiffre élevé de 26 (non compris le petit établissement arawak de Floride, qui n'est qu'une colonie d'une famille sud-américaine, mais y compris l'Eskimo, dont la place est désormais dans la famille ouralienne (14 ; 15 ; 21 ; 22)). Encore serai-je amené au cours de ce chapitre à faire des réserves sur la valeur de certains groupements, qui me paraissent avoir été constitués sans preuves linguistiques suffisantes.

Je dois cependant faire mention d'un travail récent, où P. Radin (8) propose une classification des langues nord-américaines (non compris l'Eskimo) et même d'une partie des langues de l'Amérique centrale en trois sous-groupes :

I. Sališ, Wakaš, Kutenai, Algonkin ;

II. Penutia, Klamath, Šahaptin, Uto-Aztęk, Tano, Yuki, Mixe-Zoke, Zapotęk, Kaddo, Iroquois.

III. Na-Dene, Hoka, Maya, Muskhogi.

Si intéressante que soit cette tentative de synthèse, on ne peut la considérer, dans l'état actuel de nos connaissances, que comme une « anticipation » hardie, sur la valeur réelle de laquelle l'avenir pourra seul nous fixer.

Au point de vue grammatical, on trouvera d'utiles indications pour un certain nombre de familles dans un important mémoire publié en 1908 par Uhlenbeck (**23**). Toutefois, la classification du savant linguiste hollandais comportant encore 54 groupes, alors qu'actuellement on n'en compte plus que 26, il se peut que telle particularité morphologique, signalée par lui comme commune à toutes les langues d'un groupe, ne soit plus caractéristique de la famille plus compréhensive, constituée par la fusion de plusieurs de ces groupes.

I. Famille Algonkin (**1**, I, 735-873).

Les langues de cette famille étaient parlées le long de toute la côte atlantique, depuis le Labrador jusqu'au Pamlico sound (Caroline septentrionale). Au Canada, elles atteignaient à l'Ouest les Montagnes rocheuses et au Nord la rivière Churchill. Aux États-Unis, à l'exception du pays Iroquois, tout le vaste territoire au Nord de la rivière Ohio et à l'Est du Mississippi était occupé par des Algonkin. En outre, de grandes tribus de même origine : Šeyen, Arapaho, Gros-ventres, vivaient dans les prairies occidentales ; enfin, de petites fractions avaient pénétré en Californie.

Les dialectes algonkin se répartissent en cinq groupes :

a) le groupe Blackfoot (Pieds-noirs) ;

b) le groupe Šeyen ;

c) le groupe Arapaho ;

d) le groupe central et oriental ;

e) le groupe californien.

a) Le groupe Blackfoot occupe le vaste territoire compris entre la North Saskatchewan river (Canada), le méridien 110°, le pied des Montagnes rocheuses et la frontière septentrionale de l'État de Montana. Il comprend les dialectes Piegan, Kaina ou Blood et Blackfoot ou Siksika.

b) Le groupe Šeyen, qui comprend les dialectes Šeyen et *Sutaio, se trouve dans le Sud-Est du Minnesota et le Nord-Est du Dakota méridional.

c) Le groupe Arapaho occupe la partie orientale du North Dakota, la portion adjacente du Minnesota et le Sud du Manitoba; il comprend les dialectes Gros-ventres (Atsina) et Arapaho proprement dit.

d) Le groupe central et oriental se subdivise en un sous-groupe central et en un sous-groupe méridional.

1) Le sous-groupe central comprend :

les dialectes Cree(*Kri*)-Montagnais, parlés de la rivière Churchill au Labrador, qui forment la bordure septentrionale de la famille Algonkin, auxquels se rattache le Naskapi en usage dans tout le territoire compris entre la rive orientale de la baie d'Hudson et la côte du Labrador ;

le dialecte Menomini, parlé entre les lacs Supérieur et Michigan;

les dialectes Sauk, Fox, Kikapu, sur la rive ouest du lac Michigan, et Šawni, dans le Kentucky et le Nord du Tennessee;

les dialectes Odżibwa ou Tšippewa, Algonkin, Ottawa, parlés au Nord des lacs Supérieur et Huron, et du haut Saint-Laurent; Potawatomi, en usage entre le lac Michigan et le lac Huron; les dialectes Illinois (*Ilinwa*) : *Kahokia, Kaskaskia, *Mitšigamea, *Moingwena, Peoria et *Tamaroa, parlés dans le Sud du Wisconsin, le Nord de l'Illinois et des parties de l'Iowa et du Missouri ; enfin le dialecte Miami, parlé au Sud du lac de Michigan dans le Nord-Est de l'Illinois et le Nord de l'Indiana ;

le *Natik, parlé autrefois dans le village de ce nom dans le Massachusetts;

le Delaware (Lenápe, Lenno-Lenápe, Leni-Lenápe), parlé dans tout le bassin de la Delaware dans l'Est de la Pensylvanie, le Sud-Est de l'État de New York et dans la plus grande partie de New-Jersey et Delaware (Munsi, *Unami et *Unalatštigo), auquel il faut sans doute adjoindre le *Wappinger, le *Mahikan ou *Mohikan et le Pekot, de l'Ouest du Vermont, du Massachusetts et du Connecticut, et du Sud-Est de l'État de New York.

2) Le sous-groupe oriental comprend :

le Mikmak, parlé à l'Est de l'embouchure du Saint-Laurent, jusqu'en Nouvelle-Écosse ;

l'Abnaki, parlé autrefois dans le Maine et les parties adjacentes du New-Hampshire, du Vermont et du Nouveau-Brunswick ;

le Penobskot (Pennakuk), parlé dans le Sud du New-Hampshire, l'Ouest du Vermont et le Nord du Massachusetts ;

le Passamakoddy, parlé autour de la baie de ce nom, sur la rivière Sainte-Croix et le lac Shoodic, à la limite du Maine et du Nouveau-Brunswick ;

le Malesit, très voisin du précédent, autrefois en usage sur la rivière St. John.

On range également dans le groupe central et oriental, mais sans pouvoir préciser leur position exacte à l'intérieur de ce groupe, les dialectes dits du type Massatšuset : le *Massatšuset proprement dit, le *Narraganset, le *Wampanoag, le *Nauset, le Montauk, etc..., parlés dans Long Island, l'État de Rhode Island, l'Est du Massachusetts et du Connecticut, à côté desquels il faut peut-être placer les dialectes *Nipmuk et des Indiens de la rivière Connecticut, en usage dans le Massachusetts.

Enfin, il y a un certain nombre de tribus de l'Est des États-Unis, appartenant certainement à la famille Algonkin, dont la situation linguistique reste entièrement à déterminer ; ce sont les *Nantikok, les *Konoy, les Powhatan, les *Weapemeok, les *Sekotan, etc... échelonnés du Nord au Sud le long de la côte du Maryland, de la Virginie et de la Caroline septentrionale.

e) Le groupe californien (famille Ritwan de Krœber), dont le rattachement aux précédents a été démontré par Sapir (**9** ; **9** *bis*), comprend le Wiyot (Wishoskan de Powell), parlé sur la côte de Californie depuis les montagnes Bear River jusqu'à Little river, et le Yurok (Weitspekan de Powell), parlé dans des villages du cours inférieur du Klamath (1 dialecte) et sur la côte adjacente au Sud, y compris la baie Trinidad (3 dialectes).

II. Famille *Beothuk.

La langue *Beothuk était parlée autrefois à Terre-Neuve.

III. Famille Eskimo (1, I, 967-1069 ; 19 ; 20).

Les Eskimo ou Innuit vivent sur les côtes du Groenland et le long de la côte arctique américaine, depuis le Labrador jusqu'à la rivière Copper, Alaska. Des tribus parlant Eskimo existent également dans le Nord-Est de l'Asie (Yuit) et peuplent les îles Aléoutiennes (Aleut).

Thalbitzer distingue les dialectes suivants au Groenland : sur la côte orientale, l'Angmagssalik (Âmmasalik), au niveau du 66e parallèle ; sur la côte occidentale, à l'extrême Nord, le dialecte de Smith-sound, entre le glacier Humboldt et le cap York (76°-79° de latitude), l'Upernivik (Upernawik), entre le 72e et le 74e degrés de latitude, le dialecte Umanak (Oommanaq) et de Disko-bay, parlé autour des deux grandes baies ainsi dénommées, et enfin, le dialecte en usage depuis la baie Disko jusqu'à Godthaab (70°-64° de latitude). A ces dialectes groenlandais, il adjoint les dialectes de la Terre de Baffin et du Labrador, pour constituer le groupe oriental de la famille Eskimo.

Le groupe occidental comprend les dialectes de l'île Kadiak, de Bristol bay, de l'embouchure du Yukon, de Norton et Kotzebue sound, de Point Barrow et de l'embouchure du Mackensie.

Quant au groupe central, encore mal connu, il comprend les dialectes de la rive occidentale de la baie d'Hudson, de l'île Southampton, des péninsules Melville et Boothia, et d'une partie de la Terre de Baffin.

Les Aleut sont divisés linguistiquement en Unalaska, des îles Fox et Shumagin et de l'extrémité occidentale de la péninsule d'Alaska, et en Atka, des îles Andreanof, Rat et Near.

Le Yuit comprend quatre dialectes, le Noökalit du cap East, le Aiwanat de la pointe Indian, le Wuteëlit du cap Ulakhpen, et le Eiwhuelit de l'île St.-Lawrence. Ces Eskimo asiatiques sont venus d'Amérique à une époque probablement récente.

L'Eskimo est, de toutes les langues parlées en Amérique, la seule qui ait pu jusqu'ici être rattachée à une famille linguistique de l'ancien monde. Sa parenté avec la famille ouralienne est actuellement démontrée (14 ; 15 ; 21 ; 22).

IV. Famille Hoka (3; 12).

La famille Hoka est, d'après les récents travaux de Dixon, Kroeber et Sapir, une des plus étendues de l'Amérique du Nord. Ces auteurs y classent les groupes suivants:

a) les Šasta (Sastean de Powell), qui vivaient de la rivière Rogue (Orégon) aux rivières Salmon et New (Californie) et aux sources de la rivière Sacramento, et sont apparentés aux Atšomawi et Atsugewi (Palaihnihan de Powell) du bassin de la rivière Pit;

b) les *Tšimariko, qui vivaient sur la Trinity river au Sud de l'embouchure de South fork jusqu'à la plaine de Taylor (Californie);

c) les Karok (Quoratean de Powell), qui occupent le cours moyen du Klamath (Californie);

d) les Yana, qui vivent au centre de la Californie du Nord et parlent trois dialectes;

e) les Pomo (Kulanapan de Powell), qui vivent au Nord de San Francisco (Californie) dans la vallée de la Russian river, autour du lac Clear et sur la côte, et parlent huit dialectes différents;

f) les *Esselen, autrefois installés le long de la côte de Californie, au nord des monts Santa Lucia, près de la baie Monterey;

g) les Yuma de l'Arizona, de Californie et du Mexique, que l'on répartit en trois sous-groupes:

1) les Yuma de l'Est, qui comprennent les Havasupai, les Walapai, les Tontos et les Yavapai;

2) les Yuma du centre, qui comprennent les Mohave, les Yuma proprement dits, les Marikopa, les Diegueño et les Kokopa (Cucapa);

3) les Yuma de Basse-Californie, qui comprennent les Kiliwi, les Santo Tomás et les *Kotšimi;

h) les *Salina, qui vivaient dans les deux missions de San Antonio et de San Miguel sur la côte de Californie (2 dialectes);

i) les *Tšumaš, qui habitaient les îles de Santa Barbara et la côte californienne adjacente [1];

j) les Séri, répandus en Sonora entre les parallèles 28° et 30°, le méridien 111 et le golfe de Californie;

k) les Wašo, qui vivent en Nevada et en Californie dans les environs du lac Tahoe;

l) les Tekistlatẹk (Tekisistẹk, ou Tšontal d'Oajaca, ou Tšontal d'Ecatepec), qui vivent dans l'État d'Oajaca près de l'isthme de Tehuantepec par 96° de longitude et 16° de latitude;

m) et enfin, le groupe *Koahuiltẹk, au sens large donné à ce terme par Swanton (**16**), groupe qui comprend [2]:

1) le *Koahuiltẹk proprement dit (Paikawa de Gatschet), avec quatre dialectes : le *Carrizo, le *Koahuiltẹk, le *Kotonam et le *Comecrudo, parlés sur les deux rives du bas Río Grande, auquel il faut peut-être rattacher le *Tamaulipeko, parlé autrefois par diverses tribus du Nuevo León et du Tamaulipas (Mexique);

2) le *Karankawa, parlé sur la côte du Texas;

3) le *Tonkawa, parlé dans le Sud-Ouest du Texas.

Les langues de la famille Hoka seraient caractérisées phonétiquement par le fait que les fricatives, aussi bien sourdes que sonores, qu'elles soient labiales, dentales ou palatales, sont exceptionnellement bien développées (**7**, 286).

V. Famille Iroquois (*Irokwa*).

La famille Iroquois comprend :

a) les Huron ou Wyandot, qui habitaient primitivement au Nord du Saint-Laurent, et plus tard autour du lac Simcoe (Ontario);

b) les Cinq Nations (Kayuga, Mohawk, Oneida, Onondaga, Seneka), qui vivaient dans l'État de New York;

c) les *Konestoga et *Suskehanna, établis sur la rivière Susquehanna en Pensylvanie et dans le Maryland;

1. Le Salina et le Tšumaš étaient classés auparavant en une famille indépendante, la famille Iskoman.

2. J'ai eu l'occasion de faire des réserves sur la valeur des rapprochements qui ont servi de base à la constitution de ce groupe Koahuiltẹk (**24**, 433-434), et par suite, je dois en faire sur l'incorporation de ce groupe dans la famille Hoka.

d) les Tuskarora de la Caroline septentrionale ;

e) les Tšeroki des Alleghanies méridionales.

Avec toutes réserves, Swanton suppose que les *Kori, qui vivaient aux environs du cap Lookout et du Core sound sur la côte de la Caroline du Nord, sont à rattacher à la famille Iroquois (17).

VI. Famille Kaddo.

Cette famille est divisée en trois groupes géographiques :

a) le groupe nord, qui comprend les Arikara établis sur le haut Missouri, dans le Dakota septentrional ;

b) le groupe central, qui comprend les quatre tribus Pawni de la vallée de la rivière Platte (Nebraska);

c) le groupe méridional, qui comprend les Kaddo, les Witšita et les Kitšai, autrefois établis dans l'Oklahoma, le Texas et la Louisiane.

VII. Famille Keres (1 *ter*).

Le Keres est parlé dans un certain nombre de villages du Río Grande, dans le Nouveau-Mexique. Il semble comporter deux dialectes, le dialecte des villages de l'Ouest (Sitsime ou Kawaiko): Acoma et Laguna, le dialecte des villages de l'Est (Keres proprement dit) : Cochiti, Sia, Santa Ana, San Felipe et Santo Domingo.

VIII. Famille Kiowa.

Les Kiowa ont habité autrefois le haut Yellowstone et Missouri. Actuellement, ils vivent sur le haut Arkansas et Canadian, dans le Colorado et l'Oklahoma.

IX. Famille Klamath (Lutuamian de Powell).

Le Klamath est parlé par deux tribus, les Klamath, qui habitent autour des lacs qui portent leur nom, dans la partie centrale de l'Orégon méridional, et les Modok, au Sud de ceux-ci, dans la Californie septentrionale.

X. Famille Kutenai (Kitunahan de Powell).

Les Kutenai vivent dans le Sud-Est de la Colombie britannique

et le Nord des États de Montana et d'Idaho. Ils parlent deux dialectes très voisins l'un de l'autre.

XI. FAMILLE MUSKHOGI.

Cette famille renferme un grand nombre de dialectes, autrefois parlés dans le Sud-Est des États-Unis, depuis le Mississippi jusqu'à la rivière Savannah, l'Océan Atlantique et le Golfe du Mexique, qu'on range en deux groupes :

a) Le groupe Muskhogi proprement dit, qui comprend lui-même :

1) un sous-groupe méridional, où l'on classe le Hitšiti, le Seminole, l'*Apalatši, le *Yamasi, l'Alabama (y compris le Koasati), le Tšoktaw (Chactas), avec trois dialectes, le Tšoktaw proprement dit, le Tšikasaw, et l'*Okalusa ;

2) un sous-groupe septentrional, qui comprend les Muskhogi ou Krik (Creek).

b) Le groupe Natšez, qui comprend le Natšez, l'*Avoyel et le *Taënsa.

Avec toutes réserves, Swanton (**17**) propose de rattacher à cette famille les tribus du Sud de la Floride : *Kålusa, *Ais, *Tekesta, etc., et les *Paskagula, qui vivaient sur la rivière du même nom dans le Sud du Mississippi.

XII. FAMILLE NA-DENE (**1**, I, 85-282 ; **4** *bis* ; **4** *ter* ; **13** *bis* ; **13** *ter*).

Cette famille a été constituée par Sapir (**11**), par la réunion de trois groupes de langues considérés jusqu'ici comme indépendants : le groupe Athapaskan, le groupe Haida, le groupe Tlingit.

a) Le groupe Athapaskan est de beaucoup le plus vaste de tous les groupes linguistiques nord-américains. Son domaine s'étend depuis la côte arctique jusqu'au Mexique septentrional, du Pacifique à la baie d'Hudson, au Nord, du Río Colorado à l'embouchure du Río Grande, au Sud, c'est-à-dire sur plus de 40 degrés de latitude et 75 degrés de longitude.

Phonétiquement, les langues de ce groupe sont caractérisées par des séries de sons gutturaux, très souvent continus, et de

fréquents clics et aspirations. Morphologiquement, elles se distinguent par un complexe verbal d'une grande complication, laquelle est due pour une large part à de nombreux préfixes altérés et à des changements variés de la racine destinés à indiquer le nombre et le caractère du sujet et de l'objet.

On divise le groupe en trois sous-groupes géographiques : le sous-groupe septentrional, le sous-groupe du Pacifique, le sous-groupe méridional.

1) Le sous-groupe septentrional ou sous-groupe Tinneh ou Déné comprend trois divisions :

α) La division orientale s'étend sur un territoire continu, limité à l'Ouest par les Montagnes rocheuses et le bas Mackenzie, au Sud par la ligne de partage des eaux entre la rivière Athabasca et le bas Peace, le lac Athabasca et la rivière Churchill, au Nord et à l'Est par la bande étroite et continue du territoire eskimo, en bordure de la baie d'Hudson et de l'Océan arctique. Les principales tribus que l'on classe dans cette division sont les Tatsanottine ou Yellowknives (Couteaux-jaunes),à l'Est de la rivière Yellowknife, les Thlingtšadinne ou Dogribs (Côtes-de-chien), entre les lacs Great Slave et Great Bear, les Kawtšodinneh ou Hares (Peaux-de-lièvre) et les Etšaottine ou Slavey, échelonnés du Nord au Sud sur le Mackenzie, les Tšipewayan ou Montagnais, sur la Slave river, les Tsattine ou Beavers, sur la Peace river, et enfin 500 milles au Sud, en dehors du territoire délimité plus haut, la petite tribu des Sarsi, étroitement liés avec leurs voisins algonkin, les Siksika.

β) La division du Nord-Ouest occupe l'intérieur de l'Alaska et les portions adjacentes du Canada jusqu'aux Montagnes rocheuses n'atteignant le littoral, dont elle est séparée par les Eskimo, qu'au niveau de la baie Cook et de la Copper river. Cette division comprend les Kutšin des rivières Porcupine et Tanana, du cours moyen du Yukon et du bas Mackenzie (où on les désigne souvent sous le nom de Loucheux), les Ahtena de la Copper river, et les Khotana du bas Yukon, de la rivière Koyukuk et de la baie Cook.

γ) La division du Sud-Ouest occupe la région montagneuse du Canada depuis le haut Yukon jusqu'au parallèle 51°30′, avec

comme limite orientale les Montagnes rocheuses. Elle est séparée du Pacifique par les Haida et les Tlingit, qui font partie de la même famille mais non du même groupe, par les Tsimšian et les Wakaš. Elle comprend, en commençant par le Nord, les Nahane, les Sekani, les Babine (Nataotin),sur les rives du lac de ce nom, les Carriers (Takulli), depuis le lac Stuart jusqu'à Alexandria sur la rivière Fraser, et les Tšilkotin (Tsilkotin), dans la vallée de ce nom.

2) Le sous-groupe du Pacifique comprend une petite tribu en Colombie britannique, deux autres dans l'État de Washington, et de nombreux villages répartis dans un territoire presque continu de 400 milles de longueur environ, commençant à la rivière Umpqua (Orégon) et s'étendant tout le long de la côte et des Coast Range mounts jusqu'aux sources de la rivière Eel (Californie). Ce territoire est coupé, au niveau de la rivière Klamath, par les Yurok, tribu algonkin. Ce sous-groupe comprend les dialectes suivants : le *Kwalhiokwa, dans l'État de Washington, le Umpkwa et le Kokwil (Mišikhwutmetunne), autrefois en usage sur les rivières de même nom, le *Taltuštuntude, le Šastakosta et le Tututunne, sur la rivière Rogue et ses affluents, le Tšetko, sur la rivière de même nom, dans l'Orégon, le Tolowa, sur la rivière Smith et autour de Crescent, le Hupa et le Tlelding, sur le cours inférieur de la rivière Trinity, le Whilkut, sur le haut Redwood, le Tšilula, sur le cours inférieur de cette dernière rivière et dans les collines Bald qui la séparent du Klamath, le Mattole, sur la rivière de même nom, le Sinkyone, le Lassik, le Wailaki et le Kuneste, dans la vallée de la rivière Eel, en Californie.

3) Le sous-groupe méridional occupe un vaste territoire qui embrasse une grande partie de l'Arizona et du Nouveau-Mexique, une partie du Sud de l'Utah et du Colorado, la bordure occidentale du Texas et du Kansas, et le Nord du Mexique jusqu'au 25ᵉ parallèle. Il comprend les Lipan, qui erraient autrefois depuis le Río Grande, dans le Nouveau-Mexique et au Mexique, jusqu'à la côte du golfe du Mexique, à travers le Texas, et qui sont maintenant établis dans le Nouveau-Mexique avec les Meskaleros, les Navaho, au Sud de la rivière San Juan, dans le Nord-Est de

l'Arizona et le Nord-Ouest du Nouveau-Mexique, et les Apatše (Apache), qui entourent de toutes parts les Navaho, sauf au Nord, et dont les principales tribus sont les Keretšo ou Vaqueros (divisés eux-mêmes en Meskaleros, Xikarillas, Faraones, Llaneros), les Tširikahua, les Pinaleños, les Koyoteros (avec les fractions de White Mountain et Pinal), les Arivaipa, les Xila Apatše (divisés en Xileños, Mimbreños et Mogollones), les Tontos et les Toboso (entre le 27ᵉ degré de latitude et le Río Grande del Norte).

En dehors de ces trois grands sous-groupes, il faut encore citer les Kiowa Apatše, petite tribu athapaskan qui vit au milieu des Kiowa et qui paraît venue du Nord.

b) Le groupe Haida (Skittagetan de Powell) comprend les Haida proprement dits des îles de la Reine Charlotte (Colombie britannique) et les Kaigani de l'extrémité méridionale de l'île du Prince de Galles (Alaska). On distingue dans ce groupe deux dialectes principaux, le Skidegate, parlé dans la partie centrale des îles de la Reine Charlotte, et le Masset, en usage dans la partie septentrionale de ces îles et à Howkan, Klinkwan et Kasaan (Alaska).

c) Le groupe Tlingit (Koluschan de Powell) comprend les tribus qui habitent le Sud-Est de l'Alaska, depuis la Copper river jusqu'à Dixon entrance et Portland canal. Les Tagiš, qui vivent à l'intérieur de la Colombie britannique, sur la rivière Lewis, parlent un dialecte particulier du Tlingit.

XIII. Famille Penutia (**1**, I, 283-422, 559-734, II, 1-629 ; **3** ; **13**).

Cette famille comprend un grand nombre de langues considérées, il y a peu d'années encore, comme formant des familles indépendantes. On peut la diviser provisoirement en quatre groupes : le groupe californien, le groupe de l'Orégon, le groupe Tšinuk, le groupe Tsimšian.

a) Le groupe californien comprend :

1) le Wintun (Copehan de Powell), parlé dans le Nord-Ouest de la Californie, sur le cours supérieur de la Trinity river, avec quatre principaux dialectes (septentrional, central, sud-oriental et sud-occidental);

2) le Maidu (Pujunan de Powell), parlé dans le Nord-Est de la Californie, dans un territoire qui comprend en partie la vallée du Sacramento, en partie la Sierra Nevada, avec trois dialectes, (dialecte du Nord-Ouest, dialecte du Nord-Est, dialecte du Sud ou Nišinam) ;

3) le Yokuts (Mariposan de Powell), parlé dans la portion méridionale de la grande vallée intérieure de Californie et dans les montagnes qui la bordent ;

4) le Miwok ou Miwa (Moquelumnan de Powell), parlé dans la portion centrale de la grande vallée de Californie et au Nord de la baie San Francisco, avec quatre dialectes bien caractérisés, et le *Kostano (Costanoan de Powell), parlé sur la côte de Californie depuis le Golden Gate jusqu'à Monterey, qui est la langue la plus divergente de la famille [1].

b) Le groupe de l'Orégon comprend :

1) le Takelma (Takilman de Powell), parlé sur le cours moyen de la rivière Rogue, dans l'Orégon méridional, avec deux dialectes ;

2) le groupe côtier de l'Orégon, où l'on range le Kus parlé autrefois par une petite tribu le long de la baie et de la rivière Coos, sur le cours inférieur de la rivière Coquille (Orégon), avec deux dialectes, le Hanis et le *Mulluk (*Miluk) ; le Siuslaw avec deux dialectes, l'un le Siuslaw proprement dit sur le Siuslaw, l'autre le Kuitš sur le bas Umpqua (Orégon) ; le Yakona avec deux dialectes également, le Yakina et le Alsea, parlés l'un et l'autre dans le Lincoln county (Orégon occidental) ;

3) le Kalapuya (Kalapooian de Powell), qui comprend des dialectes, anciennement assez nombreux, parlés dans la vallée Willamette (Orégon).

c) Le Tšinuk (Chinook) est parlé sur les deux rives de la rivière Columbia (Orégon), en aval de The Dalles jusqu'à la mer, et sur le cours inférieur de la rivière Willamette. Il comprend deux dialectes, le *bas Tšinuk et le haut Tšinuk.

Le haut Tšinuk comprend le Wasko et le Wišram parlés dans

1. Le Miwok et le Kostano ont été longtemps considérés comme formant une famille indépendante, appelée famille Mutsun.

la région de The Dalles, le Kathlamet et le Klakamas en usage dans la basse vallée de la Columbia et dans la vallée Willamette.

Le *bas Tšinuk comprend le *Klatsop sur la rive méridionale de la Columbia (d'Astoria en aval) et le *Tšinuk proprement dit sur la rive septentrionale (de Grays labor en aval) et dans la baie Shoalwater.

d) Le Tsimšian (Chimmesyan de Powell) est parlé le long de la côte septentrionale de la Colombie britannique. Il comporte trois dialectes : le Tsimšian proprement dit, parlé sur la rivière Skeena et dans les îles au Sud, le Niska de la rivière Nass, le Gyitkšan sur le cours supérieur de la Skeena.

XIV. Famille Šahaptin.

Cette famille comprend un certain nombre de tribus qui vivent dans le Sud-Ouest de l'Idaho, le Sud-Est de l'État de Washington et le Nord-Est de l'Orégon. Les plus connues sont les Klikitat, les Nez Percés, les Palus, les Topiniš, les Umatilla, les Wallawalla, les Warm Springs ou Tenino et les Yakima.

XV. Famille Sališ.

Le Sališ est parlé dans le Nord de l'État de Washington et de l'Idaho, l'Ouest du Montana, une petite étendue de la côte de l'Orégon, la partie sud-est de Vancouver, et le Sud de la Colombie britannique. En plus, il existe deux groupes isolés : l'un en pays kutenai, l'autre sur la côte de la Colombie britannique, sur les rives de Dean inlet, de Burke channel et de la rivière Bellacoola.

Les divers dialectes qu'il comporte sont classés géographiquement de la façon suivante :

a) Dialectes de l'intérieur : Lilluet, Ntlakyapamuk, Šuswap, Okinagan en Colombie britannique, Flathead dans Washington, Idaho et Montana, Skitswiš ou Cœur d'Alène dans le Nord d'Idaho, groupe Columbia dans l'Ouest de Washington ;

b) Dialectes de la côte : Bellakula, Komoks, Kowitšan, Skihwamiš, Songiš, Niskwalli, Twana, Tšehalis, Tillamuk.

XVI. Famille Siou (*Syu*) (1, I, 875-965).

Les Algonkin occupaient les grandes forêts de l'Est, les Sioux les grandes plaines à buffles, c'est-à dire, en grandes lignes, toute la région comprise entre le Mississippi et les Montagnes rocheuses, du Canada à la rivière Arkansas. Un second groupe vivait dans le centre de la Caroline septentrionale, de la Caroline méridionale et de la Virginie. Enfin, il y avait deux petites tribus isolées dans le Mississippi : les Biloksi sur la côte du golfe du Mexique, et les *Ofo, sur la rivière Yazoo.

On classe les divers dialectes Sioux en sept groupes :

a) le groupe Dakota-Assiniboin, qui comprend les Mdewakanton, les Wahpekute, les Sisseton, les Wahpeton, les Yankton, les Yanktonai, les Teton et les Assiniboin ;

b) le groupe Dhegiha, qui comprend les Omaha, les Ponka, les Kwapaw, les Osage (*Ozaž*) et les Kansa ;

c) le groupe Tšiwere, qui comprend les Iowa, les Oto, les Missouri et les Winnebago ;

d) le groupe Mandan ;

e) le groupe Hidatsa, qui comprend les Hidatsa et les Crow ;

f) le groupe Biloksi, qui comprend les Biloksi et les *Ofo ;

g) le groupe oriental, qui comprend les *Tutelo aux sources du Dan, les Katawba de la Caroline du Sud et quelques dialectes éteints et douteux.

XVII. Famille Tano.

Cette famille comprend des dialectes parlés dans des villages de la vallée du Río Grande, que l'on répartit en trois groupes :

a) le Tiwa, parlé dans les villages de Taos, Picuris, Sandia, Isleta, Isleta del Sur, auquel on rattache le *Piro de Senecu et Socorro del Sur ;

b) le *Towa, parlé à Jemez et dans l'ancien village de Pecos ;

c) le Tewa, parlé à San Juan, Santa Clara, San Ildefonso, Nambe, Pojoaque, Tesuque et Hano .

XVIII. Famille *Timukua.

Le *Timukua était parlé par un groupe de tribus dans la Floride septentrionale.

XIX. Famille Tšimakum.

Cette famille comprend les *Tšimakum, qui vivaient dans la partie occidentale de l'État de Washington dans la péninsule située entre le Hood canal et Port Townsend, et les Kwileut, établis sur la côte de l'État de Washington, au Sud du cap Flattery, dont une sous-tribu porte le nom de Hoh.

XX. Famille *Tunika (**18**).

Cette famille, revue récemment et considérablement élargie par Swanton, comprend [1] :

a) le *Tunika parlé sur les deux rives du Mississippi, près de l'embouchure du Yazoo (dialectes *Tunika, *Griga ou *Gri, *Yazu, *Koroa et *Tiu) ;

b) l' *Atakapa du Sud-Ouest de la Louisiane (dialectes *Atakapa, *Akokisa, *Bidai et *Opelusa ?) ;

c) le Tšitimaša du Sud de la Louisiane, auquel il faut adjoindre le *Waša et le *Tšawaša, parlés sur le bayou La Fourche, près de l'embouchure du Mississippi.

XXI. Famille Uto-Aztẹk.

Sous ce nom, Brinton, en s'appuyant sur les recherches de Buschmann, a réuni, dès 1891, trois groupes de langues, considérés pendant longtemps comme indépendants : Šošon, Pima-Sonora, Nahuatl (**2**, 119-134) ; mais la légitimité de ce groupement n'a été vraiment bien établie que grâce aux récents travaux de Kroeber et de Sapir (**6** ; **10**).

a) Le groupe Šošon est un des plus vastes de l'Amérique du Nord. Il occupe le Sud-Ouest du Montana, l'Idaho en entier, au Sud de 45° 30′ de latitude environ, le Sud-Est de l'Orégon, au Sud des monts Blue, l'Ouest et le centre du Wyoming et du Colorado, une partie du Nord et l'Est du Nouveau-Mexique, le Nord-Ouest du Texas, l'Utah, une partie de l'Arizona septentrional, le

1. Je dois faire toutes réserves sur la légitimité de la constitution de ce groupe, l'étude que j'ai faite de l'Atakapa et du Tšitimaša ne m'ayant pas révélé d'affinités nettes entre ces deux idiomes (**24**, 437-442) et les arguments donnés par Swanton ne me paraissant pas décisifs.

Nevada (à l'exception de la petite enclave des Wašo), un petit territoire dans le Nord-Est de la Californie à l'Est de la Cordillère et un vaste espace dans l'Est du même État au Sud du 38e degré environ, comprenant les cours supérieurs de quelques affluents du San Joaquín, les pentes méridionales de la Sierra et la côte du Pacifique entre les 33e et 34e degrés de latitude.

En outre, suivant une tradition, les Šošon auraient occupé autrefois le Sud-Ouest de l'État d'Alberta et le Nord-Ouest du Montana.

Le groupe Šošon est divisé en quatre sous-groupes :

1) le sous-groupe du Plateau, qui comprend les tribus des vastes espaces arides du Great Basin, c'est-à-dire les Šošoni-Komãš (Gosiute, Komãš, Šošoni et Šikaviyam de Californie), les Ute-Tšemehuevi (Ute septentrionaux et méridionaux, Panamint, Paiute du Sud, Tšemehuevi et Kawaiisu de Californie) et les Mono-Paviotso ou Mono-Bannok (Bannok, Paiute du Nord ou Paviotso, Mono et Šošon de l'Orégon oriental);

2) le sous-groupe de la Californie méridionale, qui comprend le Serrano (avec divers dialectes : Möhineyam, Gitanemuk, etc.), le Luiseño-Kahuilla (avec les dialectes San Luiseño, San Juaneño, Agua Caliente et Kahuilla) et le Gabrieliño (avec les dialectes Gabrieliño et Fernandino) ;

3) le sous-groupe de la rivière Kern (Californie), qui est représenté par les Tübatulabal ;

4) le sous-groupe Pueblo [1], qui comprend le Hopi ou Moki de Tusayan (Arizona).

b) Le groupe Pima-Sonora occupe ou occupait le Sud de l'Arizona et le Nord-Ouest du Mexique, c'est-à-dire la plus grande partie des États de Sonora, Chihuahua, Sinaloa et Durango, et des portions de Jalisco et de Zacatecas.

Ce groupe comprend :

1) les Pima altos (Upper Pima), installés dans les vallées Gila

1. Les Indiens désignés sous le nom de Pueblos ne forment pas une unité linguistique : ce terme correspond à une unité archéologique, qui comprend, outre les Hopi, les tribus des familles Tano, Keres et Zuñi.

et Salt du Sud de l'Arizona, auxquels se rattachent étroitement les Pápago, qui vivaient originellement dans la région située au Sud et au Sud-Est de la rivière Gila, spécialementau Sud de Tucson (Arizona) et s'étendentà l'Ouest et au Sud-Ouest, à travers le désert désolé connu sous le nom de Papaguería, jusqu'en Sonora (Mexique), les *Sobaipuri, qui vivaient dans le bassin des ríos San Pedro et Santa Cruz, entre les 110ᵉ et 111ᵉ degrés de longitude et sur le río Gila entre l'embouchure du San Pedro et les ruines de Casa grande et peut-être aussi dans le Sud de l'Arizona, et les *Potlapigua de la région de Babiste à la limite nord-est du territoire ópata ;

2) les Pima bajos (Névome, Lower Pima), qui s'étendent à l'Est et à l'Ouest, le long du bas et du moyen Yaqui, avec deux petites colonies, les Pima de Bamoa, au Sud du Mayo sur le bas Sinaloa, et les Tepehuan, dans l'État de Durango, au Sud des Tarahumare, principalement sur les pentes orientales de la Sierra Madre entre les 24ᵉ et 27ᵉ degrés de latitude ;

3) les Ópata (Teguima), établis aux sources du Yaqui et de l'Hermosillo, dont on connaît deux dialectes, l'Eudeve (Heve ou Dohema) et le Xova (Jobal ou Ova) ;

4) les Kahita (Yaqui, Cinaloa, Sinaloa), établis principalement sur le cours inférieur et moyen du Yaqui, du Mayo et du río del Fuerte, depuis le golfe de Californie jusqu'à la Sierra, dont les principaux dialectes sont le Yaki, le Mayo, le Tehueko et le Vakoregüe ;

5) les *Tepahue, installés dans les montagnes, près de la division supérieure du río Mayo, dans le Sud de Sonora, près de la frontière septentrionale du territoire kahita ;

6) les *Zoe, qui occupaient une petite région près dela frontière orientale du territoire kahita, à la source du río del Fuerte, auxquels on rattache les *Baimena (Baitrena) installés immédiatement au Sud ;

7) les *Nio, qui vivaient dans le Nord de Sinaloa, dans la région où se trouve la ville actuelle qui porte leur nom ;

8) les Tarahumare, qui habitent les régions montagneuses d'une partie de Sonora, de Chihuahua et de Durango et parlent les dialectes Varohio, Guazápare, Patšera et Tubar ;

9) les *Kontšo, sur l'affluent de gauche du Río Grande qui porte ce nom [1];

10) les *Laguneros ou *Irritila, installés autour des lacs des plateaux de Mapimi, principalement autour du lac de Parras ou San Pedro [1];

11) les *Akaxee et les tribus apparentées *Xixime, *Tebaka, *Sabaibo, dans les montagnes des États de Durango et de Sinaloa;

12) les *Zakatẹk, qui habitent l'État de Zacatecas et une partie de l'État de Durango, notamment dans le bassin du río Navas [1];

13) les Huitšol (Guichola), qui habitent à l'Est des précédents dans la Sierra Madre (Nord-Ouest de l'État de Jalisco), et qui sont sans doute les descendants des *Guatšitšile (Cuachichile);

14) les Kora (Chora, Chota, Nayarita), qui vivent dans la Sierra de Nayarit et sur le río Jesus María (État de Jalisco);

15) les Tepekano, établis dans la sierra du Jalisco septentrional, en particulier dans le village d'Askeltán, apparentés aux Teule ou Teul-Tšitšimẹk des environs de Teul.

c) Le groupe Nahuatl comprend :

1) le sous-groupe Nahuatl (Aztẹk ou Aztèque proprement dit ou Mexicano), qui occupe une plus ou moins grande étendue des États de Tabasco, Vera Cruz, Chiapas, Puebla, Tlaxcala, Guerrero, México, Hidalgo, Morelos, Michoacán, Colima, San Luis, Sinaloa, Durango, Zacatecas, Tepic, Jalisco, avec de petites colonies dans l'État d'Oajaca, dans les districts de Pochutla (1 *bis*), de Tuxtepec, de Teotitlan et de Silacayoapam;

2) le Pipil, qui comporte plusieurs dialectes : le Pipil parlé le long de la côte du Pacifique dans le Sud-Ouest du Guatémala, depuis Escuintla jusqu'à la rivière Lempa (Salvador), avec interruption par le groupe Xinka, le Pipil du haut Motagua dans l'Est du Guatémala, auquel on rattache l'*Alaguilak parlé autrefois dans la région de San Cristobal sur la rive gauche du haut Motapan, et, avec doute, le Pipil du Honduras (au Sud de Trujillo);

3) le *Nikarao (Olomega ou Nikiran), parlé autrefois entre

1. Je classe ces trois idiomes dans le groupe Pima-Sonora plutôt que dans le groupe Nahuatl uniquement pour des raisons géographiques.

la côte du Pacifique et le lac Nicaragua, dans les îles de ce lac, sur la rive méridionale du golfe de Fonseca, et enfin dans le Costa Rica, au Nord du golfe de Nicoya ;

4) le Tlaskaltẹk, parlé aux environs d'Izalco (Salvador) ;

5) le *Sigua (Xicagua, Segua, Shelaba, Chicagua, Chichagua), parlé par une petite colonie mexicaine établie sur la côte nord du Costa Rica, dans les environs du lac Chiriquí ;

6) le *Kazkan, parlé sur la rive droite du Río Grande de Santiago [1].

XXII. Famille Waiilatpu.

Le Waiilatpu était parlé par deux tribus de l'Orégon, les *Kayuse, établis aux sources des rivières Walla Walla, Umatilla et Grande Ronde, et les Molala, qui vivaient entre le Mt. Hood et le Mt. Scott.

XXIII. Famille Wakaš (1, I, 423-557).

Cette famille comprend deux groupes de dialectes :

a) le Nutka, parlé sur la côte occidentale de l'île Vancouver et aux environs du cap Flattery ;

b) le Kwakiutl, qu'on divise en trois sous-groupes :

1) le sous-groupe septentrional, qui comprend le dialecte parlé autour de la baie Gardner et du détroit Douglas ;

2) le sous-groupe central, aux environs de Milbank sound et de la baie Rivers ;

3) le sous-groupe méridional, avec quatre dialectes, dont seul le Kwakiutl de Vancouver a été bien étudié.

XXIV. Famille Yuki.

Cette famille comprend quatre dialectes profondément différenciés, le Yuki proprement dit parlé dans Round valley sur la rivière Eel, Californie, le Yuki de la côte occidentale, entre Ten Mile river et Rockport ou Usal, le Hutšnom de la South Eel river, et le Wappo aux sources de la Russian river.

1. Il est possible que cette langue appartienne au groupe Pima-Sonora.

XXV. Famille Yutši (Uchean de Powell).

Les Yutši vivaient autrefois sur la rivière Savannah (Géorgie ils habitent maintenant avec les Krik dans l'Oklahoma.

XXVI. Famille Zuñi.

Le Zuñi, parlé dans le village de ce nom dans le Nouveau-Mexique, était autrefois en usage dans sept villages de la même région.

INDEX BIBLIOGRAPHIQUE.

1. Boas (Franz). *Handbook of american indian languages*. Bureau of american ethnology, Bulletin 40. Washington, t. I, 1911, t. II, 1922.

1 *bis*. Boas (Franz). *El dialecto mexicano de Pochutla, Oaxaca*. International Journal of american linguistics. New York, t. I, 1917-1920, p. 9-44.

1 *ter*. Boas (Franz). *A Keresan text*. International Journal of american linguistics. New York, t. II, 1921-1923, p. 171-180.

2. Brinton (Daniel G.). *The american race*. New York, 1891.

3. Dixon (Roland B.) et Krœber (A. L.). *Linguistic families of California*. University of California Publications in american archaeology and ethnology. Berkeley, t. XVI, 1919-1920, p. 103-112.

4. Goddard (Pliny Earle). *The present conditions of our knowledge of North American languages*. American anthropologist. Lancaster, new series, t. XVI, 1914, p. 555-592.

4 *bis*. Goddard (Pliny Earle). *Wailaki texts*. International Journal of american linguistics. New York, t. II, 1921-1923, p. 77-135.

4 *ter*. Haeberlin (Hermann K.). *Notes on the composition of the verbal complex in Haida*. International Journal of american linguistics. New York, t. II, 1921-1923, p. 159-162.

5. *Handbook of american Indians North of Mexico* (edited by Frederick Webb Hodge). Bureau of american ethnology, Bulletin 30, 2 vol. Washington, 1907-1910.

6. Krœber (A. L.). *Shoshonean dialects of California*. University of California Publications in american archaeology and ethnology. Berkeley, t. IV, 1906-1907, p. 65-166.

7. Krœber (A. L.). *Serian, Tequistlatecan and Hokan*. University of California Publications in american archaeology and ethnology. Berkeley, t. XI, 1911-1916, p. 279-290.

8. Radin (Paul). *The genetic relationship of the North American indian languages*. University of California Publications in american archaeology and ethnology. Berkeley, t. XIV, 1918-1919, p. 489-502.

9. Sapir (Edward). *Wiyot and Yurok, Algonkin languages of California*. American anthropologist. Lancaster, new series, t. XV, 1913, p. 617-646.

9 *bis*. SAPIR (Edward). *The Algonkin affinity of Yurok and Wiyot kinship terms*. Journal de la Société des Américanistes de Paris, nouvelle série, t. XV, 1923, p. 36-74.

10. SAPIR (Edward). *Southern Paiute and Nahuatl. A study in Uto-Aztekan*. Journal de la Société des Américanistes de Paris, nouvelle série, t. X, 1913, p. 379-425 ; t. XI, 1914-1919, p. 443-488.

11. SAPIR (Edward). *The Na-Dene languages, a preliminary report*. American anthropologist. Lancaster, new series, t. XVII, 1915, p. 534-558.

12. SAPIR (Edward). *A supplementary note on Salinan and Washo*. International Journal of american linguistics. New York, t. II, 1921-1923, p. 68-72.

13. SAPIR (Edward). *A characteristic Penutian form of stem*. International Journal of american linguistics. New York, t. II, 1921-1923, p. 58-67.

13 *bis*. SAPIR (Edward). *A type of Athabaskan relative*. International Journal of american linguistics. New York, t. II, 1921-1923, p. 136-142.

13 *ter*. SAPIR (Edward). *The phonetics of Haida*. International Journal of american linguistics. New York, t. II, 1921-1923, p. 143-158.

14. SAUVAGEOT (Aurélien). *Sur certaines similitudes de l'Eskimo et des langues ouraliennes*. Bulletin de la Société de linguistique de Paris ; Communication faite à la séance du 17 novembre 1923 (sous presse).

15. SAUVAGEOT (Aurélien). *Eskimo et Ouralien*. Journal de la Société des Américanistes de Paris, nouvelle série, t. XVI, 1924.

16. SWANTON (John R.). *Linguistic position of the tribes of southern Texas and northeastern Mexico*. American anthropologist. Lancaster, new series, t. XVIII, 1915, p. 17-40.

17. SWANTON (John R.). *Unclassified languages of the Southeast*. International Journal of american linguistics. New York, t. I, 1917-1920, p. 47-49.

18. SWANTON (John R.). *The Tunica language*. International Journal of american linguistics. New York, t. II, 1921-1923, p. 1-39.

19. THALBITZER (William). *The Aleutian language compared with Greenlandic ; a manuscript by Rasmus Rask, dating from 1820, now in the Royal Library at Copenhagen*. International Journal of american linguistics. New York, t. II, 1921-1923, p. 40-57.

20. THALBITZER (W.). *The Ammasalik Eskimo*, part II : *Language and folklore*. Meddelelser om Grönland. Copenhague, t. XL, 1923, p. 113-564.

21. UHLENBECK (C. C.). *Uralische Anklänge in den Eskimosprachen*. Zeitschrift der deutschen morgenländischen Gesellschaft. Leipzig, t. LIX, 1905, p. 757-765.

22. UHLENBECK (C.C.). *Zur Eskimogrammatik*. Zeitschrift der deutschen morgenländischen Gesellschaft. Leipzig, t. LX, 1906, p. 112-114.

23. UHLENBECK (C. C.). *Die einheimischen Sprachen Nord-Amerikas bis zum Rio Grande*. Anthropos. Vienne, t. III, 1908, p. 773-799.

24. VILLIERS DU TERRAGE (M. de) et RIVET (P.). *Les Indiens du Texas et les expéditions françaises de 1720 et 1721 à la « Baie Saint-Bernard »*. Journal de la Société des Américanistes de Paris, nouvelle série, t. XI, 1914-1919, p. 403-442.

II. LANGUES DE L'AMÉRIQUE CENTRALE.

Pour la classification des langues de l'Amérique centrale, je me suis principalement inspiré de la révision que l'on doit à C. Thomas et J. R. Swanton (**6**) et de la belle monographie, récemment publiée par W. Lehmann (**4**). Je n'ai pas cru toutefois pouvoir accepter tous les rapprochements proposés par ce dernier, les preuves linguistiques, sur lesquelles il s'est appuyé, ne m'ayant pas paru toujours suffisamment convaincantes. De nouvelles recherches me semblent nécessaires pour établir d'une façon définitive le bien fondé des hardies conceptions du savant ethnographe allemand.

Les familles énumérées dans ce chapitre sont exclusivement celles qui sont spéciales à l'Amérique centrale. Les autres ont été étudiées dans le chapitre précédent (Uto-Aztẹk, Na-Dene, Hoka), ou seront étudiées avec les langues de l'Amérique du Sud (Tšibtša, Karib).

I. Famille Amusgo.

Walther Lehmann classe à part l'Amusgo (Amishgo, Amuchco), qui était jusqu'ici rangé dans la famille Mixtẹk. La tribu qui parlait cette langue occupait en pays mixtẹk une étroite bande de territoire, à partir du Pacifique, entre les 16ᵉ et 17ᵉ degrés de latitude (États de Guerrero et d'Oajaca).

II. Famille Kuikatẹk (**1**).

Le Kuikatẹk, que Mechling avait cru pouvoir classer dans la famille Mixtẹk (**5**), est considéré par Walther Lehmann comme un idiome indépendant. La tribu qui parle cette langue habite le district de Cuicatlán (Oajaca).

III. Famille Kuitlatęk.

Le Kuitlatęk, qu'on classait jusqu'ici comme dialectes Nahuatl, est considéré par W. Lehmann comme un idiome indépendant. Les Kuitlatęk (ou Teko, ou Popoloka du Michoacán) habitaient autrefois une vaste région dont la capitale était Mexcaltepec sur la côte du Pacifique. Leur langue était parlée plus récemment à Ajuchitlán, à San Cristóbal, à Poliutla et à Atoyac (État de Guerrero).

IV. Famille Lenka.

Les Lenka occupaient autrefois une vaste région dans le Honduras central et occidental, atteignant le Pacifique à travers la partie du Salvador comprise entre la rivière Lempa et la baie de Fonseca.

Le Lenka comprend les dialectes Guaxikero, *Intibukat, Opatoro et Similaton (Honduras central), Tšilanga, *Kakaguatike et *Guataxiagua, désignés parfois sous le nom de Tšontal (Nord-Est du Salvador).

V. Famille Maya (7).

La famille Maya est la plus importante des familles spéciales à l'Amérique centrale (le Nahuatl étant désormais rattaché à la famille nord-américaine Uto-Aztęk).

La meilleure classification que l'on ait des multiples dialectes Maya est celle de Stoll ; c'est elle que je reproduis avec de très légères modifications.

La famille Maya comprend deux groupes : le groupe Huastęk et le groupe Maya proprement dit.

a) Dans le groupe Huastęk, on classe le Huastęk, parlé le long du golfe du Mexique depuis Vera-Cruz jusqu'à San Luis Potosí, et à quelque distance à l'intérieur de l'État de Tamaulipas, et son dialecte, le Tšikomuseltęk, parlé dans le Sud-Est du Chiapas.

b) Le groupe Maya se divise lui-même en deux sous-groupes : le sous-groupe Tzental-Maya et le sous-groupe Pokontší-Kitšé-Mam.

1) Le sous-groupe Tzental-Maya comprend :

α) la division Maya, qui occupe les États du Yucatan et de Campeche, la partie orientale des Etats de Tabasco et de Chiapas, le Nord du Guatémala et le Honduras britannique presque en entier, et où l'on compte quatre dialectes : le Maya proprement dit (Mayathan), parlé dans toute la péninsule du Yucatan, le Nord du Honduras britannique, la partie occidentale de l'État de Tabasco et le Nord-Ouest du Guatémala ; le Lakandon, en usage dans la région montagneuse du haut Usumacinta dans le Nord-Ouest du Guatémala et l'Est de l'État de Chiapas ; l'Itza ou Peten, parlé dans le Guatémala septentrional (à l'exception du petit territoire lakandon à l'Ouest) et la région adjacente du Honduras anglais ; enfin, le Mopan, parlé dans le Sud du Honduras britannique et la portion adjacente du Guatémala ;

β) la division Tzotzil, qui comprend le Tšontal de Tabasco ; le Tzental (Tzendal), parlé dans l'État de Chiapas ; le Tzotzil (Zotzil), parlé dans l'État de Chiapas, qui correspond sans doute à la langue des anciens *Quelene ; le Tšañabal (Tojolabal), parlé par une petite tribu à l'extrémité sud-est du Chiapas, près de la frontière du Guatémala ; le Tšol, parlé dans le Nord-Est de l'État de Chiapas sur la rive gauche du río Usumacinta, dans l'Est du Guatémala, le Sud du Honduras britannique et le Nord-Ouest du Honduras, sur le río Polochic et le Golfo dulce, ainsi que sur le bas Motagua, auquel se rattachent les dialectes suivants : le Tšuxe (Chuhe), en usage de Nenton à San Sebastian près de la frontière occidentale du Guatémala, le Xakaltęk, parlé aux environs de Jacaltenango, près de la frontière nord-occidentale du Guatémala, et le Motozintlęk, parlé aux environs de Motozintla à l'extrême Sud-Est du Chiapas ; le Tšortí, parlé sur la lisière orientale du Guatémala (province de Chiquimula) et dans la partie adjacente du Honduras (Copan), et enfin le *Subinha, lexicographiquement très près du Tšañabal et du Tzental [1].

2) Le sous-groupe Pokontší-Kitšé-Mam comprend trois divisions :

1. L'emplacement de cette tribu n'a pu être déterminé avec certitude. Thomas et Swanton supposent qu'elle habitait dans l'Est de l'État de Chiapas.

α) la division Pokontší, qui comprend le Pokontší, parlé aux sources du Cahabon ; le Kektší (K'aktchí), parlé par une importante tribu du Guatémala central sur les deux rives du haut Cahabon jusqu'à la rivière Chixoy à l'Ouest ; le Pokomám (Pokam), parlé dans le Sud-Est du Guatémala, depuis le Río Grande ou haut Motagua jusqu'à la frontière du Guatémala et du Salvador ;

β) la division Kitšé, qui comprend le Kitšé proprement dit, parlé dans une vaste région du Guatémala central aux sources du río Motagua et le long de la rive occidentale du lac Atitlan jusqu'au Pacifique ; le Kaktšikel, parlé entre le lac Atitlan et les environs de la ville de Guatémala, dont le Pupuluka de St. Mary, près d'Antigua, n'est qu'un dialecte ; le Tzutuhil (Zutuhil), parlé dans un petit territoire près de la rive méridionale du lac Atitlan ; l'Uspantęk, parlé par une petite tribu du Guatémala central dans la grande boucle du río Chixoy (río Negro) ;

γ) la division Mam, qui comprend le Mam proprement dit (Zaklohpakap), parlé dans le Guatémala occidental depuis Soconusco environ jusqu'au Pacifique ; le Ixil, parlé dans les villages de Nebaj, Cotzal et Chajul (Guatémala central) ; l'Aguakatęk I de Stoll, parlé dans une petite région autour de Aguacatan et de Huehuetenango, et (avec doute) le *Atšis, parlé autrefois dans l'Est du Guatémala.

VI. Famille Mískito-Sumo-Matagalpa.

W. Lehmann rapporte à une souche commune les langues des groupes Mískito, Úlūa ou Sumo et Matagalpa, qu'on considérait antérieurement comme indépendantes.

a) Le Mískito (Muskito, Moskito) est parlé sur la côte atlantique du Nicaragua et du Honduras, depuis le río Patuca au Nord, jusqu'à un point intermédiaire entre les ríos Bluefields et San Juan au Sud.

Les principales tribus qu'on classe dans ce groupe sont les Mískito proprement dits, établis le long du littoral ; les Táŭirra, à l'intérieur des terres entre le río Coco et le Prinzapolca ; les Mām, au Nord du bas Coco jusqu'au Patuca ; les Wankí, sur le río Coco, depuis son embouchure jusqu'à Bocay.

b) Le groupe Úlũa ou Sumo (Simoo, Smoo) occupe une grande partie du Nicaragua oriental et le Sud du Honduras. Il est divisé en trois sous-groupes :

1) le sous-groupe Úlũa, au Nord-Est du lac de Nicaragua ;

2) le sous-groupe Sumo-Táũaxka, qui a des représentants sur le Putuk, mais occupe surtout le bassin du rio Coco et le cours supérieur des rivières qui se jettent dans l'Atlantique entre ce fleuve et le Prinzapolca ;

3) le sous-groupe Yósko, qui occupe le bassin des ríos Tuma, Hamaca et Līsaũči.

c) Le Matagalpa ou Tšontal du Nicaragua (appelé parfois aussi Popoluka de Matagalpa) est parlé dans une région ayant pour centre la ville de Matagalpa et comprenant le district de Matagalpa presque entier ainsi qu'une partie des districts de Segovia et de Chontales ; deux de ses dialectes sont, en outre, en usage dans les villages de Cacaopera et de Lislique, à l'extrémité nord-est du Salvador.

VII. Famille Mixe-Zoke.

Dans cette famille, Walther Lehmann range :

le Mixe (Mizè), parlé dans les districts de Choapan, Juchitán, Yautepec, Villa-alta et Tehuantepec (Oajaca), auquel se rattachent trois dialectes de l'État de Vera-Cruz : le Popoluka de Oluta et de Texistepec et le dialecte de Sayula ;

le Zoke, parlé dans une partie des États de Chiapas, Tabasco et Oajaca, avec un dialecte, le Tapixulapan, parlé dans les villages de Tapijulapa, Ocsolotan et Puzcatan, de l'État de Tabasco ;

le *Tapatšultẹk, parlé autrefois dans l'extrême Sud-Est de l'État de Chiapas et au Sud du pays mam, notamment dans le village de Tapachula ;

l'*Aguakatẹk II de Stoll, parlé dans les villages de Aguacatan et de Chalchitan, aux sources du río Negro (Guatémala) ;

le Huave (Huabi, Juave, Guavi, Wabi), autrefois parlé par une tribu installée dans la région qui entoure la Laguna superior et la Laguna inferior, dans les districts de Juchitán et de Tehuantepec (Oajaca).

VIII. Famille Mixtẹk.

Les Mixtẹk habitent un territoire qui s'étend de la côte du Pacifique à la région montagneuse de l'intérieur dans les États de Guerrero et de Puebla, et surtout dans l'Oajaca occidental. La division de cette famille en Mixteca alta et Mixteca baja est purement géographique.

IX. Famille *Olive.

Les *Olive étaient une petite tribu qui vivait à l'extrémité méridionale de l'État de Tamaulipas.

X. Famille Otomí (Tšitšimẹk de Peñafiel et Buelna).

On classe ordinairement dans cette famille une série de langues parlées par des peuplades qui occupent un vaste territoire du Mexique central, s'étendant depuis México jusqu'au 22e degré de latitude nord et, de l'Est à l'Ouest, sur 4 degrés environ de longitude, comprenant une partie des États de México, de San Luis Potosí, de Vera-Cruz, de Tlaxcala et de Michoacán, les États de Guanajuato et de Hidalgo presque entiers, tout l'État de Querétaro, avec deux îlots dans l'État de Puebla. Ces langues sont : l'Otomí proprement dit (Hiă-hiŭ), parlé dans la plus grande partie de l'État de Querétaro, dans une portion de celui de Guanajuato, spécialement dans les districts de Celaya, Cortazar, Iturbide, Chamacuero de Comonfort et San Luis de la Paz, dans diverses localités des États de San Luis Potosí, Hidalgo, Michoacán, Puebla, Vera-Cruz et Tlaxcala, et enfin dans le district fédéral, notamment à Tacubaya, Guadalupe, Hidalgo, Tlalpan et Xochimilco ; le *Serrano, parlé autrefois dans le territoire de Sierra gorda (formé avec des parties des États de Querétaro, Guanajuato et San Luis Potosí) ; le *Meko ou *Xonaz ou *Tonaz, de la même région, peut-être identique au précédent, en usage notamment dans les missions de San Luis de la Paz (Guanajuato) et de San José Vizarron (Querétaro) ; le Tepehua, parlé à Huehuetla, dans le district de Tenango (État de Hidalgo), à Huayacocotla, canton de Chicontepec (État de Vera-Cruz) et à

Tlaxco (État de Puebla) ; le Pame (ou Tšitšimęk [1]), parlé à la mission de Cerro prieto (État de México), dans quelques villages et missions de San Luis Potosí, de Querétaro et de Guanajuato (avec trois dialectes) ; le Mazahua, parlé au Sud-Ouest de l'État de México, où il est encore en usage dans le district de Ixtlahuaca, et le Pirinda (ou Matlaltzinco), parlé au Sud du même État et à Charo (Michoacán).

A la famille Otomí, ainsi comprise, W. Lehmann propose de rattacher :

a) le groupe Mazatęk de Mechling (**5**), qui comprend :

1) le Trike, parlé dans les hautes montagnes des districts de Tlaxiaco et de Juxtlahuaca (Oajaca) ;

2) le Tšotšo (ou Tšutšon), où l'on range le Tšotšo proprement dit, ou Popoloko d'Oajaca, parlé dans les districts de Coixtlahuaca et de Teposcolula (Oajaca), et le Popoloko de l'État de Puebla ;

3) le Mazatęk, parlé dans les districts de Teloloapan (État de Guerrero), de Tuxtepec, de Cuicatlán et de Teotitlan (Oajaca), avec ses trois dialectes, le Mazatęk, le Ixkatęk et le *Guatekimame ;

b) le Tšiapanęk, considéré jusqu'ici comme indépendant, qui comprend :

1) le Tšiapanęk proprement dit, parlé dans les villages d'Alcala, de Chiapa et de Suchiapa (État de Chiapas) ;

2) le Mangue (Tšolutęk), parlé au Nord du pays des Subtiaba, le long du golfe de Fonseca jusqu'au Honduras ;

3) le *Diriá, autrefois en usage dans le territoire compris entre l'extrémité supérieure du lac Nicaragua, la rivière Tipitapa, la moitié sud du lac Managua et le Pacifique ;

4) l'Orotina, parlé dans le Nord-Ouest du Costa Rica depuis la rive méridionale du lac Nicaragua au Nord, le long des deux rives du golfe de Nicoya, jusqu'au Pacifique à l'Ouest [2].

1. En fait, le nom de Tšitšimęk n'a aucun sens linguistique précis ; il a été donné à toutes les tribus errantes et guerrières des États de Zacatecas, Aguas calientes, Querétaro et San Luis Potosí, à quelque famille linguistique qu'elles appartiennent : Zakatęk, Teule, Kazkan, Guatšitšile, Irritila, Pame et Otomí.

2. Le Mangue, le Diriá et l'Orotina étaient considérés auparavant comme formant une famille indépendante, la famille Tšorotęk.

XI. Famille Paya.

Les Paya vivent au Honduras, dans le triangle compris entre le río Tinto, la mer et le río Wawks ou Segovia.

XII. Famille Subtiaba.

Cette famille comprend d'après W. Lehmann :

a) le Tlappanẹk-Yopi de l'État de Guerrero ;

b) le Subtiaba (Maribio d'Oviedo, Maribi de Berendt, Nagrando de Squier), parlé dans la région comprise entre l'extrémité septentrionale du lac Managua et le Pacifique, c'est-à-dire dans la plus grande partie du district de León (Nicaragua) ;

c) le *Maribitšikoa, parlé autrefois sur le río Guatahiguala, dont W. Lehmann identifie le nom avec celui du village de Guatajiagua du Salvador.

XIII. Famille Tarask (Michoacano).

Le Tarask est parlé dans la plus grande partie de l'État de Michoacán.

XIV. Famille Totonak.

Les Totonak habitaient sur la côte orientale du Mexique un territoire qui commence au 21e degré de latitude et correspond à la partie septentrionale des États de Vera-Cruz et de Puebla et à l'extrémité orientale de l'État d'Hidalgo.

Le Totonak comprenait plusieurs dialectes, le Tetikilhati, parlé dans la cordillère, le Tšakahuaxti, dans les villages de Xalpan et de Pantepec, le Tatimolo, dans le village de Naolingo, et l'Ipapana, dans les missions des Pères Augustins.

XV. Famille Tšinantẹk.

Le Tšinantẹk est parlé dans les districts de Choapan, Tuxtepec et Ixtlan de l'État d'Oajaca, et à la limite occidentale de l'État de Vera-Cruz.

XVI. FAMILLE *WAÏKURI.

Cette famille comprend le *Waïkuri et le *Perikú, parlés autrefois dans l'extrémité méridionale de la Basse-Californie.

XVII. FAMILLE *XANAMBRE (Janambre).

Cette famille comprend le *Pisone et le *Xanambre, parlés dans le Sud-Ouest de Tamaulipas, de la vallée de Purísima au río Blanco.

XVIII. FAMILLE XIKAKE (Jicaque) (3).

Le Xikake est parlé par une tribu habitant le Honduras septentrional, du río Ulva à l'Ouest, au río Negro à l'Est. On en connaît trois dialectes, le Xikake de Yoro, le Xikake de Palmar, et le Leán y Mulia.

XIX. FAMILLE XINKA (Jinca, Sinca).

Le Xinca, appelé parfois Popoloka du Guatémala, est parlé par un petit groupe vivant dans le Sud-Est du Guatémala, le long de la côte, depuis le río Michatoyat jusqu'à la frontière du Salvador. On connaît trois dialectes de cette langue : le Sinakantan, le Xupiltepęk et le Xutiapa, parlés dans les villages de ce nom, auxquels il faut peut-être ajouter, comme dialecte très différencié, le *Pupuluka de Conguaco, dans l'extrême Sud du Guatémala.

XX. FAMILLE ZAPOTĘK (2).

La famille Zapotęk comprend :

a) le Zapotęk proprement dit, parlé dans un vaste territoire au centre et au Sud-Est de l'État d'Oajaca (à peu près la moitié de l'État) ;

b) le Soltęk, parlé dans le district de Zimatlán (Oajaca) ;

c) le Tšatino, parlé dans les districts de Juquila, Juxtlahuaca, Teojomulco et Yautepec (Oajaca) ;

d) le Papabuko, parlé à Elotepec et dans quelques autres villages de l'État d'Oajaca.

INDEX BIBLIOGRAPHIQUE.

1. ADÁN (Elfego). *Los Cuicatecos actuales*. Anales del Museo nacional de arqueología, historia y etnografía. México, 4e série, t. I, 1922, p. 137-154.

2. BOAS (Franz). *Notes on the Chatino language of Mexico*. American anthropologist. Lancaster, new series, t. XV, 1913, p. 78-86.

3. CONZEMIUS (Edward). *The Jicaques of Honduras*. International Journal of american linguistics. New York, t. II, 1921-1923, p. 163-170.

4. LEHMANN (Walther). *Zentral-Amerika*. I Teil. *Die Sprachen Zentral-Amerikas*. Berlin, 1920, 2 vol.

5. MECHLING (William H.). *The indian linguistic stocks of Oaxaca, Mexico*. American anthropologist. Lancaster, new series, t. XIV, 1912, p. 643-682.

6. THOMAS (Cyrus) et SWANTON (John R.). *Indian languages of Mexico and Central America and their geographical distribution*. Bureau of american ethnology, Bulletin 44. Washington, 1911.

7. TOZZER (Alfred M.). *A Maya grammar with bibliography and appraisement of the works noted*. Papers of the Peabody Museum of american archaeology and ethnology, Harvard University. Cambridge, t. IX, 1921.

III. LANGUES DE L'AMÉRIQUE DU SUD ET DES ANTILLES.

Je réunis dans un seul chapitre les langues de l'Amérique méridionale et des Antilles, parce que les populations indiennes des Antilles se rattachent à deux familles linguistiques sud-américaines: la famille Arawak et la famille Karib. Ce rattachement n'augmente donc pas d'une unité le nombre des familles énumérées ci-après, qui s'élève à 77.

Ce chiffre considérable montre à quel point les études linguistiques sud-américaines sont peu avancées. A vrai dire, le travail de classement est à peine ébauché, et si certains grands groupements sont déjà constitués, il y a une foule d'idiomes, dont les affinités exactes sont à déterminer et que notre ignorance oblige à considérer pour l'instant comme indépendants.

Dans l'exposé qui suit, je me suis surtout servi des deux essais de classification les plus récents dus à Ehrenreich (**16**) et à Chamberlain (**9**). Pour éviter des répétitions constantes, je n'ai pas mentionné ces deux travaux fondamentaux parmi les références qui se rapportent à chaque famille, mais le lecteur devra toujours s'y reporter, car c'est là qu'il trouvera les renseignements bibliographiques essentiels, que je ne pouvais songer à reproduire ici. Les références que je donne pour chaque famille se rapportent en effet exclusivement aux travaux parus depuis la publication des études de Ehrenreich et de Chamberlain, quelquefois à des travaux que ces auteurs ont oublié de citer ou ignorés.

Lorsqu'une famille a été l'objet d'une révision récente ou qu'un travail à elle consacré permet d'atteindre de proche en proche tout ou partie des publications antérieures, je n'indique que cette seule référence, même s'il s'agit d'une étude beaucoup

moins importante que les études précédentes, auxquelles elle renvoie. Les références que je donne n'ont d'autre but que de fournir aux chercheurs la *clef* de la bibliographie complète de chaque famille.

I. Famille Alakaluf (9 *bis* ; 39) (Alikuluf de Chamberlain).

Lehmann-Nitsche, à qui l'on doit la révision la plus récente de cette famille, y range les tribus suivantes, énumérées en allant du Nord au Sud :

les *Tšono, qui habitaient l'archipel chilien du même nom ;

les *Kaukahue, qui habitaient les îles de ce nom, sans doute l'archipel actuel de Wellington ;

les *Letšeyel et les *Yekinahue ou *Yekinahuere, qui vivaient au Nord du détroit de Magellan ;

les *Enoo ou *Pešerä des îles du détroit de Magellan, auxquels se rattachent les *Adwipliin de l'île Londonderry ;

les Alikulip (Alukulup, Alakaluf), qui s'appellent eux-mêmes Hékaïné, qui vivent entre la partie occidentale du canal Beagle et le détroit de Magellan.

II. Famille *Al'entiak (74).

La famille *Al'entiak comprend :

les *Al'entiak ou *Huarpe (Guarpe), qui habitaient les environs des grands lacs de Huanacache, et s'étendaient probablement jusqu'aux pentes occidentales de la sierra de Córdoba, et, dans la direction du Sud, jusqu'à la partie septentrionale des provinces de San Luis et de Mendoza ;

les *Mil'kayak, qui habitaient la province de Cuyo, c'est-à-dire les provinces argentines modernes de Mendoza, San Juan et San Luis de la Punta.

III. Famille Amueša (Lorenzo de Chamberlain) (71 ; 78).

Les Amueša (Amueixa) vivent sur le río Colorado, affluent du Chanchamayo, sur le Paucartambo, et surtout dans le bassin du Palcazu, affluent du Pachitea. Suivant Tello, l'Amueša serait une langue arawak.

IV. Famille Araukan (Aukanian de Brinton ; - Che de Lehmann-Nitsche) (2 *bis* ; 8 *ter* ; 38).

Les Araukan (Auka) occupaient et occupent encore en partie la région chilienne comprise entre le Pacifique et la cordillère des Andes, depuis Copiapó au Nord jusqu'à Chiloé au Sud, entre les 27ᵉ et 43ᵉ degrés de latitude environ, les pentes orientales de la cordillère depuis la province de San Juan et la lagune de Guanacache au Nord jusqu'au Limay et à la lagune de Nahuelhuapi au Sud, le gouvernement de Neuquén, une partie de celui du Río Negro et de la pampa argentine (où ils ont été désignés sous le nom de Molutše par Falkner et de Pueltše par Camaño), jusqu'aux environs de Buenos Aires à l'Est.

Les Araukan sont divisés en plusieurs groupes, qui parlent ou parlaient des dialectes peu différenciés d'une même langue, le Maputše. Ce sont les *Pikuntu ou *Pikuntše (gens du Nord), entre Coquimbo et le 35ᵉ parallèle et jusqu'à Mendoza à l'Est ; les Pehuentše (gens des forêts de pins), depuis le 35ᵉ degré de latitude jusqu'à Valdivia et sur les pentes orientales de la Cordillère à la source du Neuquén ; les Kunko ou Huil'itše (gens du Sud), depuis le río de Valdivia au Nord jusqu'à l'archipel de Chiloé et le lac de Nahuelhuapi au Sud, auxquels appartiennent les Manzanero ou Molutše (gens de l'Ouest), établis sur les deux rives du Limay et dans les environs des lacs Lacar et Nahuelhuapi ; les *Velitše ou *Tšilote de l'archipel de Chiloé ; les Taluhet ou Talutše, à l'Est des Pikuntše et du río Salado jusqu'aux lagunes de Guanacache au Nord, et par petits groupes, sur les rives des ríos 2°, 3° et 4° dans la province de Córdoba ; les *Diuihet ou *Diuitše, à l'Est des Pehuentše, entre les ríos Colorado, Atuel et Salado ; les Leuvutše, sur les deux rives du río Negro en aval de Neuquén ; les Rankel, entre le río 5° et le río Colorado, aux sources du Chalileo.

V. Famille Arawak (Maipure de Gilii ; Nu-Aruak de von den Steinen ; Arowak de Ehrenreich) (2 *ter* ; 14 ; 17 ; 17 *bis*, 23-52, 62-76, 78 ; 27 ; 33 ; 46 *ter* ; 52 *bis* ; 54 ; 64 ; 65 ; 69, 211-218 ; 72 *bis* ; 78 ; 80).

La famille Arawak est une des plus importantes, sinon la plus importante, des familles linguistiques de l'Amérique du Sud. C'est du moins celle qui a le domaine le plus vaste. Elle compte en effet des représentants depuis l'extrémité méridionale de la Floride au Nord jusqu'au Paraguay septentrional au Sud, depuis l'Océan Pacifique (côte péruvienne) à l'Ouest jusqu'à l'embouchure de l'Amazone à l'Est.

Dans l'énumération des multiples tribus qui la composent, je suivrai l'ordre géographique ; j'exposerai ensuite comment une partie d'entre elles se groupent d'après leurs affinités linguistiques.

Au moment de la découverte, l'Arawak était parlé dans toutes les Antilles, grandes et petites ; mais dans les petites Antilles, il s'était produit un phénomène étrange. Ces îles ayant été envahies, peu de temps avant l'arrivée des Espagnols, par les Karib de Guyane (voir Famille Karib), la population arawak avait été dans certains cas, comme à la Trinité, refoulée dans les montagnes de l'intérieur; mais presque partout, les hommes ayant été massacrés par les envahisseurs et les femmes étant devenues les épouses de ceux-ci, il s'était créé une symbiose linguistique très curieuse : une langue d'origine arawak, réservée aux femmes et aux enfants en bas-âge, coexista avec une langue nettement karib parlée par les hommes seuls. Cette situation paradoxale ne fut pas transitoire ; elle s'est perpétuée jusqu'à nos jours chez les Karib de la Dominique et on peut en retrouver des traces même chez les Karib du Honduras.

Les *Arawak des Antilles sont désignés par les anciens auteurs sous différents noms. Dans toutes les petites Antilles, de la Trinité à Porto Rico, on les appelait *Alluag (Allouague), *Ineri, *Inyeri, *Igneri, *Eyeri, *Kabre ; à Aruba, Bonaire et Curaçâo, *Kaketío ; à Haïti, *Taino ; dans les Bahamas, *Lukayan. De Cuba, les Arawak gagnèrent le Sud-Ouest de la Floride, où ils avaient un village en territoire kalusa.

Sur la terre ferme, les Arawak occupaient autrefois très probablement la plus grande partie du bas pays vénézuélien et toute la côte atlantique entre l'embouchure de l'Orénoque et l'Ama-

zone ; ils en furent chassés en partie, peu de temps avant la conquête, par les Karib; toutefois, il reste encore plusieurs tribus arawak dans cette région. En Guyane franco-brésilienne, ce sont les Marawan(Palikur, Okawan, Rukuan), établis sur le Couripi, sur ses deux affluents, l'Ouassa et le Rocaoua, et sur le bas Oyapok ; les Arawak proprement dits (Aruak, Aroaqui, Arawaak, Aroaco, Arawack), qui s'appellent eux-mêmes Lukkunu, et vivent en Guyane britannique entre les rivières Corentyn et Pomeroon et sur l'Aruka, affluent occidental de la Barima ; les Taruma (avec une branche éteinte, les *Parauien), sur le versant guyanais des monts Acarahy[1] ; les Atoraí (Ataroi, Aturrai, Atorad, Atorradi, Atorayo, Aturati), qui occupaient, il y a cinquante ans encore, les plaines entre le Rupununi et le Cuduwini (Cuyuwimi), affluents de l'Essequibo, les monts Carawaina et les sources du Tacutú, affluent du río Branco, mais qui ont été absorbés par les Wapišána ; les Mapidian (Maopityan, Moonpidenne, Pidian), sous-tribu des Atoraí, qui vivaient encore en 1884 sur le versant brésilien des monts Acarahy, à la source de l'Apiniwau ou Curucuri, région qu'ils ont abandonnée pour gagner le Sud de la Guyane anglaise à travers le pays taruma ; les Wapišána (Wapisiana, Wapityan, Wabijana, Mapisiana, Mauixiana, Uabixana, Uapixana, Uapichana, Vapeschana), qui, du bassin du Tacutú et de ses affluents septentrionaux, le Mahú et le Surumú, ont envahi, à l'Ouest, les affluents septentrionaux du bas Uraricuera, notamment le Majuri, au Sud, les vastes savanes des deux rives du río Branco, à l'Est, la région montagneuse qui fait le partage des eaux de l'Essequibo et du río Branco, le bassin du Rupununi et toutes les plaines guyanaises au Sud de cette rivière, après avoir absorbé successivement les Paravilhana (tribu Karib) et les Atoraí.

Sur la côte du Vénézuéla, vivait au moment de la découverte l'importante tribu des *Kaketío, apparentée à la tribu du même nom qui peuplait les îles Aruba, Bonaire et Curaçāo, et qui,

1. Ces Indiens devaient vivre auparavant sur l'Amazone, car ce sont eux qui aidèrent à construire Fortalezza da Barra, l'actuelle Manaos.

comme celle-ci, doit être classée parmi les Arawak. Elle occupait le littoral depuis les rives du lac de Maracaibo, vers 10°30′, à l'Ouest, jusqu'au delà de l'embouchure du Yaracuy à l'Est, et s'infiltrait par le bassin de ce dernier fleuve, dans la direction du Sud-Ouest, sur les pentes orientales de la cordillère andine dans les États de Cojedes, Portuguesa et Zamora, jusqu'à l'Ele, au Sud, où les Atšagua lui donnaient le nom de Támud. On rattache aux Kaketío les *Axagua, des sources du Tocuyo.

A l'Ouest des Kaketío, vivent encore les Goaxiro, avec la sous-tribu des Kosina, dans la péninsule de Goajira, et leurs proches parents, les Parauxano, sur les bords de la partie méridionale du golfe de Maracaibo et de la partie septentrionale du lac de même nom, ainsi que dans l'île Zapara, qui représentent les anciens *Toa et *Zapara, habitants des îles ainsi dénommées.

La région où les Arawak forment actuellement le bloc le plus compact englobe le bassin de l'Orénoque et des affluents septentrionaux de l'Amazone, ríos Negro, Yapurá et Putumayo.

Dans le bassin de l'Orénoque, on rencontre les tribus suivantes : les Guinaú (identiques sans doute aux anciens *Guaniare), aux sources mêmes du Caura ; les Amorúa, entre le bas Meta et le bas Vichada, descendants probables des *Maipure signalés par les anciens auteurs sur l'Orénoque, vers 5° de latitude ; les Piapóko ou Dzásẹ, sur le bas Guaviare, et les Mitua, sur la même rivière en aval de son confluent avec l'Ariari, qui sont, selon toute vraisemblance, les descendants des *Kabre ou *Kaberre, qui l'occupaient autrefois depuis son embouchure jusqu'à l'Ariari ; les Yavitero (Paraene, Parene, Pareni) des sources de l'Atabapo ; les Mawakwá, sur le Mavaka, affluent de gauche du haut Orénoque ; les Atšagua, sur l'Ele, le Casanare, le Guachiría, l'Amuturi, le Casimena, le Cusiana, l'Upia, le haut Meta, le Muco, et entre le Meta et l'Ariari, auxquels on peut, semble-t-il, rattacher les *Tekua (Tegua, Tergua) du bassin du Lengupa, affluent de l'Upia ; les Amarizama (Amarizano), sur les rives de la lagune et du río Vua et du río Aguas blancas, affluents du Guaviare ; les Tšukuna (Tšukunẹ), sur le Manacacia et le Vichada, et vraisemblablement les *Guayupe (Guaipe), qui occupaient les rives

de l'Ariari, l'espace compris entre lui et le Guayabero, et le cours du Guaviare.

Les tribus des affluents septentrionaux de l'Amazone sont : les Baniwa (Baniva, Baníba), du Guainia et de l'Atabapo ; les Adzáneni (Tatú-tapuyo), du haut Guainía, de ses affluents méridionaux et des sources du Curary, affluent de l'Içana ; les Baré, maîtres du río Negro presque tout entier dans la première moitié du siècle dernier, actuellement cantonnés sur son cours supérieur, le bas Cassiquiare et ses affluents, particulièrement le Pasimoni ; les Masáka, sur le Cassiquiare et son affluent de gauche, le Siápa ; les Pauišana, sur la rive droite du río Branco entre les affluents Catrimani et Mocajahy ; les Yabaána, sur le haut Pasimoni et sur le Marauyá, affluent de gauche du moyen río Negro ; les Mandauáka, aux sources du río Cauaburý, affluent de gauche du moyen río Negro, et antérieurement sur le Cassiquiare, le bas Siápa et le haut Pasimoni ; les Uarekéna (Uaréka), sur le Xié et le Guainía, affluents de droite du río Negro ; les Kuatí, les Tapiíra, les Payoarini, les Ipéka, les Siusí, les Káua, les Huhúteni, les Katapolítani et les Karútana (Karúzana), échelonnés d'amont en aval dans le bassin de l'Içana et de ses affluents, quelquefois désignés sous le nom général de Izanéni ; les Íyäine et les Tariána, sur le cours moyen du Caiarý Uaupés ; les Kariay, sur le río Negro près du río Blanco ; les Manáo, près de l'embouchure et sur le cours inférieur du río Negro ; les Kauyarí, sur le haut Apaporís ; les Matapý-tapuyo et les Yukúna, sur le Miritíparaná, affluent de gauche du Yapurá, immédiatement en amont de l'Apaporis ; les Uiriná, sur le río Marari, affluent du Marauyá ; les Uainumá, sur le Yapurá dans les forêts entre l'Upí et le Cauinarý, et leurs proches parents, les Mariaté (Muriaté) ; les Yumána, entre l'Içá et le Yapurá, notamment sur les ríos Joami et Puré ; les Kauišána, sur le bas Yapurá et le Tonantins vers 2° 30' de latitude ; et les Passé, actuellement sur le bas Içá, vers 2° 30', autrefois dans toute la vaste région comprise entre le río Negro et le Putumayo.

Au Sud de l'Amazone, les tribus arawak s'égaillent davantage qu'au Nord du grand fleuve. La tribu la plus orientale est celle

des Aruá (Aroá) de l'île de Marajó. Aux sources du Xingú, se trouvent les Waurá sur la rive gauche du Batový vers 12°, les Kustenaú sur la rive droite de la même rivière vers 12°20′, les Mehinakú et les Yaulapíti sur la rive gauche du Kulisehú, respectivement vers 12°34′ et entre 12°14′ et 12°18′.

Dans le bassin du Madeira également, les Arawak sont surtout cantonnés sur le cours supérieur du fleuve. Les tribus qui y ont été signalées sont les Pama, sur la rive gauche en amont du Maparana ; les Moxo, sur les deux rives du Mamoré entre 14° et 15° de latitude ; les Baure, sur le río Baures entre 15°40′ et 14°30′, et leurs proches parents, les *Mutšoxeone, de la mission de El Carmen; les *Paikoneka, aux sources du Baures et du río Paragua ou Serre, affluent de gauche du Guaporé ; les *Paunaka, au Sud-Est des précédents, aux sources du Baures ; les *Saraveka, sur le río Verde, affluent de gauche du Guaporé, entre 14° et 16° de latitude ; les Paressí (Ariti), sur la rive droite du Guaporé, aux sources du Tapajoz et du Paraguay, dans la cordillère qui porte leur nom ; les Inapari ou Maško-Piro, sur la rive gauche du Madre de Dios entre le Tacuatimanú et l'Amigo ; les Huatšipairi, sur la rive droite du Cosñipata et du Pilcopata, d'où ils s'étendent jusqu'au Marcapata ; les Sirineiri (Maško ou Moeno), entre le Pilcopata et le Colorado, et en amont du Pilcopata, sur la rive droite du Mánu, et enfin les Lapatšu (Apolista), dans la région montagneuse qui, à l'Est d'Apolobamba, sépare le bassin du Tuichi de celui du Huanay.

La majeure partie de la population du Purús est arawak. Ce sont les *Purupurú, qui, au XVIIe siècle, occupaient les rives du fleuve depuis l'embouchure jusqu'à 50 lieues dans l'intérieur, représentés actuellement par les Pammari (Paumari), dans les îles et lagunes du moyen Purús entre l'embouchure du Jacaré et Hyatanaham, et les Yuberi du bas Tapauá ; les Pamana, sur l'Ituxy et le Mucuim ; les Yamamadi (Kapaná, Kapinamari, Kólö), dans les forêts situées entre le Purús et le Juruá, dans un territoire limité par le Mamoria-mirim, le Paûini, affluents du Purús, et la rive droite du Chiruan, affluent du Juruá ; les Ipuriná (Kángütü, Kángite, Kangiti), qui occupent le Purús et la rive droite de ce

fleuve depuis le Sepatynim jusqu'au fleuve Hyacú, les rives de l'Aquiry jusqu'au parallèle 9°45', et de l'Ituxy, aux sources duquel ils portent le nom de Kašarari ; les Uainamari, sur la rive gauche du Purús en amont de l'Hyacú, dans l'intérieur des terres ; les Maniteneri, sur le Purús entre 69° et 70°45' de longitude et aux sources de l'Aquiry ; les Kanamari, sur l'Hyacú, aux sources de l'Irariapé, affluent de gauche de l'Aquiry, et de l'Ituxy ; les Tšontakiro, sur l'Aracá ; les Kušitínéri (Kušitšinéri, Kuxixeneri), sur le Curumahá, et les Katiăná, aux sources de cette rivière.

Dans le bassin du Juruá, il y a seulement trois tribus arawak : les Marawá, sur la rive gauche du bas fleuve jusqu'au Jutahý à l'Ouest ; les Araua, sur le bas Chiué et le bas Chiruan, affluents de la rive droite ; et les Kolina (Kulina, Kulino, Kurina), autrefois établis sur la même rive depuis le Mararý et le haut Tapauá à l'Est, jusqu'au Gregorio à l'Ouest, et entre le Tarauacá et l'Envirá au Sud, actuellement disparus de la zone comprise entre le Chiruan et le Tarauacá.

A la source du Jutahý, vivent les Kuniba (Kunibo), et, sur sa rive gauche, les Uarayku (Wareku, Araykú), qui s'étendent, par les sources du Jandiatuba, jusqu'à la rive droite du Javarý.

Enfin, sur l'Ucayali, on rencontre les deux importantes tribus des Tšontakiro et des Kampa. Les Tšontakiro (Piro), dont j'ai signalé plus haut une peuplade dans le bassin du Purús, vivent sur la ligne de partage des eaux entre ce fleuve et l'Ucayali (aux sources du Sepehua et du Cujar), sur le haut Ucayali et le cours inférieur de son affluent, l'Urubamba. Quant aux Kampa, leurs multiples tribus : Anti, Kamatika, Kimbiri, Pangoa, Katongo, Kirinairi, Matši ganga, Pukapakuri, Tampa, Ugunitširi, Ungonino, occupent les bassins du Tambo, du Perené, de l'Ené, de l'Apurimac, de l'Urubamba et du Yavero.

Vers l'Ouest et vers le Sud, les Arawak ont largement débordé le bassin amazonien.

Le groupe méridional est constitué par les Guaná. Ces Indiens, qui s'appellent eux-mêmes Tšané, habitaient, au moment de la découverte, le triangle compris entre le Salado et le Paraguay, où vivent actuellement des tribus des familles Maskoi et Samuku ;

de cette importante tribu, qui a émigré vers le Nord-Nord-Est, il ne subsiste que de faibles restes, les Kinikinao à l'Ouest du bourg d'Albuquerque, les Tereno (Terena), les Guaná proprement dits (qui ont également une petite colonie dans la banlieue même de Cuyabá) et les Layaná dans les environs de Miranda. A cette fraction arawak, se rattachent, sans doute, les Tšané de l'Itiyuro et du Parapití, tribu guaranisée, dont Nordenskiöld a démontré l'origine arawak (**48**, 145-146).

Le groupe occidental est représenté par les Uru-Pukina. Les Uru (Uro, Otšozuma), actuellement réduits à de petits groupes disséminés au milieu des Aymará dans la haute Bolivie, le long du Desaguadero, dans l'île Panza du lac Poopó et dans le petit village de Chipaya au Nord de la lagune de Coipasa (sous-tribu des Tšipaya), occupaient, à une époque ancienne, toute l'immense région des hauts plateaux andins qui s'étend du Nord du lac Titicaca à la frontière argentine, et le littoral péruvien depuis Arequipa au Nord jusqu'à Cobija au Sud, point où ils se confondaient peut-être avec les *Tšango, qui peuplaient la côte chilienne jusqu'à Huasco. La langue uru n'est autre que le *Pukina, signalé par les anciens auteurs comme une des « lenguas generales » de l'ancien royaume du Pérou.

Je crois qu'on peut considérer comme un dialecte arawak très altéré le Tikuna parlé par les Indiens de même nom, qui vivaient autrefois sur l'Amazone depuis un point situé en aval de S. Paulo de Olivença jusqu'au delà de Loreto, et qu'on rencontre actuellement sur le bas Jandiatuba, entre le bas Javarý et l'Amazone, et entre l'Ambiyacu et l'Atacuary.

Je rattache enfin à la famille Arawak un groupe de langues classé habituellement sous le nom de famille linguistique Takana. J'ai montré en effet que, si ces langues ont une grammaire qui présente de grandes analogies morphologiques avec celle de la famille Pano, leur vocabulaire est en grande partie d'origine arawak (**15**).

Le territoire occupé par le groupe Takana englobe le cours supérieur des ríos Tahuamanú et Abuná et peut-être du río Aquiry, le cours du Madre de Dios entre 67° et 68°35′ de lon-

gitude et celui de ses affluents, notamment du Tambopata et du Heath, le cours du Beni entre 12° et 15° environ de latitude et celui de ses affluents, du Madidi et du Tuichi en particulier.

Les principales tribus takana sont les Araona et les Kavina (Kaviña), clans exogamiques d'une même peuplade, sur le Madre de Dios entre 67° et 68°35′ de longitude environ, sur le Manuripi et aux sources du Tahuamanú et de l'Abuná, auxquels il faut sans doute rattacher les Kapetšene de l'Aquiry (entre 9°45′ et 10°45′ de latitude) ; les Mabenaro des affluents méridionaux du haut Manuripi ; les Tiatinagua (Baguaja, Baguajairi, Mohino, Tšuntšo, Echoja, Guarayo, Huanayo, Quinaqui), sur le Tambopata et le Heath; les Toromona, entre 12° et 13° de latitude dans le territoire compris entre le Beni, le Madidi et le Madre de Dios ; les Guakanahua (appelés aussi Guarayo), sur le Madidi et l'Undumo, affluents de gauche du Beni, dont une peuplade, installée sur le Madidi, porte le nom de Tšāma ; les Kavineño, qui sont les Kavina christianisés de la mission de Cavinas, fondée d'abord sur le Madidi, actuellement sur la rive droite du Beni ; les Takana proprement dits, sur le Tuichi et au Nord de ce fleuve, qui parlent deux dialectes : le dialecte d'Isiamas ou Ydiama et le dialecte de Tumupasa ou Marakáni ; les Maropa, primitivement sur le Beni, puis réunis dans la mission de Reyes, auxquels se rattachent les *Tširiba (Tširigua) des environs de Reyes et de San Borja ; les *Sapibokona et les *Guarisa de Reyes [1].

Il n'existe pas de grammaire comparée des multiples dialectes de la famille Arawak ; toutefois, on peut ranger un certain nombre d'entre eux, d'après leurs affinités linguistiques, dans quelques groupes assez solidement établis.

Ces groupes sont :

a) le groupe nord-amazonien, qui comprend à peu près tous les dialectes de l'Orénoque et des affluents septentrionaux de l'Amazone, le Goaxiro de la péninsule de Goajira, le Yaulapíti, le Mehinakú, le Kustenaú et le Waurá du Xingú, le Paressí et le Savareka de Bolivie ;

1. En outre, deux tribus pano, les Arasa et les Atsahuaka, parlent, en plus de leur langue primitive, le Takana.

b) le groupe préandin, qui comprend l'Ipuriná, le Piro-Tšontakiro-Kuniba-Kušitinéri, le Kanamari, le Maniteneri, l'Inapari, le Kampa et le Palikur-Marawan ;

c) le groupe bolivien, qui comprend le Baure et son co-dialecte le Mutšoxeone, le Moxo, le Paikoneka et le Paunaka ;

d) le groupe Araua, qui comprend le Pama, le Pamana, le Pammari, le Purupurú, le Yuberi, l'Araua, le Yamamadi, le Kulina et, comme rameau divergent, le Guaná-Tereno-Layaná du haut Paraguay ;

e) le groupe guyanais, qui comprend le Taruma, l'Atoraí, le Mapidian et le Wapišána ;

f) le groupe Uru-Pukina ;

g) le groupe Takana.

Le centre de dispersion des Arawak paraît être la région vénézuélo-brésilienne, correspondant aux bassins de l'Orénoque et du río Negro. La date relative et l'ordre des diverses migrations parties de ce centre ne peuvent être fixées pour l'instant. On peut affirmer cependant que la migration Uru-Pukina doit être une des plus anciennes ; en effet, la langue parlée par ces tribus semble dériver de l'Arawak commun, avant toute différenciation dialectale. Au contraire, le groupe préandin présente des affinités manifestes avec le groupe nord-amazonien, dont il s'est détaché probablement à une date tardive. De tous les groupes arawak, ce sont les groupes Araua et Takana qui sont le plus différenciés. Leur caractère aberrant peut s'expliquer par le fait que les envahisseurs arawak auraient imposé leur langue à des peuplades de parler différent.

L'histoire des migrations arawak est certainement très compliquée ; c'est ainsi qu'en Bolivie, par exemple, se trouvent actuellement juxtaposées des peuplades appartenant manifestement à trois groupes différents. Seule, une étude linguistique comparée sérieuse, qui reste en grande partie à faire, permettra de reconstituer cette histoire dans ses grandes lignes.

VI. Famille *Arda.

Les *Arda habitaient entre le haut Nanay, affluent de gauche du haut Amazone, et le haut Mazán, affluent de droite du bas Napo.

VII. FAMILLE *ATAKAMA (**6**, 58-67).

Les *Atakama vivaient dans la région d'Atacama du 19ᵉ au 24ᵉ degré de latitude. Pour des raisons archéologiques, Boman pense que leur domaine s'étendait en outre autrefois à la puna de Jujuy et à tout le territoire compris entre la puna argentine et le Pacifique. Les descendants des Atakama s'appelaient eux-mêmes Likan-antai et leur langue portait le nom de Kunza.

VIII. FAMILLE *ATAL'AN (**81**, 45-61).

Je donne ce nom à une famille linguistique où je range une série de tribus disparues de la côte équatorienne, qui, ethnographiquement, paraissent apparentées : les *Manta, établis entre l'embouchure du Chone et l'île Salango ; les *Huankavilka, qui vivaient dans la région de Guayaquil ; les *Puna, qui occupaient l'île de ce nom, et enfin les *Tumbez, qui dominaient le littoral depuis le río Naranjal jusqu'au Sud du río Tumbez. Le nom « atal'an » désignait la langue parlée dans les plaines côtières de l'ancien royaume de Quito.

IX. FAMILLE AUAKÉ (**33**).

Les Auaké, réduits à une très petite tribu, peut-être à une famille, vivent aux sources du Parauá, affluent de gauche du Caróni.

X. FAMILLE AYMARÁ.

Actuellement, l'Aymará est parlé, au Pérou, dans deux provinces du département de Puno : Cercado de Puno et Chucuito, c'est-à-dire sur toute la rive sud-ouest du lac Titicaca (la ville de Puno marquant assez nettement la limite entre cette langue et le Kitšua), dans les départements d'Arequipa et de Moquegua, par les groupes indiens qui n'ont pas encore adopté l'espagnol, et enfin dans le département de Lima, où un dialecte aymará, le Kauki, parlé autrefois dans le district de Pampas de la province de Yauyos, à Huantan, à Cachui et à Aquicha, est encore en usage dans les villages de Tupe, Huaquis et Laraos et, d'après J. C. Tello, dans quelques villages des provinces de Huarochirí et de Canta. En Bolivie, le domaine aymará comprend tout le

département de La Paz (provinces de Muñecas, Omasuyu, Cercado de La Paz, Pacajes, Sicasica, Inquisivi, Larecaja, Chulumani et Apolobamba) et une partie du département d'Oruro, englobant, à l'Est, une partie de la province de Chayanta et, au Sud, la région du lac Poopó.

Le domaine aymará actuel ne correspond plus qu'à une partie du domaine ancien ; il s'est en effet considérablement rétréci, avant et depuis la découverte, surtout au bénéfice du Kitšua. En 1795, la province, aujourd'hui chilienne, d'Arica était encore habitée par 12.870 Aymará ; en 1581, l'élément dominant de la province de Lipes était aymará ; il en était de même anciennement dans la province de Chichas, alors qu'actuellement on ne parle plus que le Kitšua dans ces deux provinces. Dans la direction du Nord, l'extension ancienne des Aymará n'a pas été moins considérable ; dépassant largement les rives septentrionales du Titicaca, deux de leurs tribus, les *Kana et les *Kantši, qui parlaient certainement encore leur langue primitive au commencement du XVIIe siècle, remontaient le long de la vallée de Vilcanota, plus au Nord que Cuzcō, jusqu'à l'Urubamba. La langue *Tšumbivilka, parlée encore en 1586 dans la province de ce nom sur les ríos Santo Tomas et Velille, affluents de gauche de l'Apurimac, était, suivant toute probabilité, un dialecte aymará. En 1795, la province des Aimaraes, située sur le haut Pachachaca, affluent de gauche de l'Apurimac, était encore habitée par 10.782 Aymará. La province d'Andahuaylas était peuplée des mêmes Indiens. Plus à l'Ouest, l'Aymará était encore en usage, en 1586, conjointement avec le Kitšua, dans le Sud du département d'Ayacucho, dans l'ancienne province de Vilcas Huaman sur le cours supérieur du río Pampas (pays des anciens *Tšanka). Enfin, la toponymie permet, semble-t-il, d'inclure dans le domaine aymará une partie de la province de Huancavelica, et même des provinces de Lima, Tarma et Huarochirí, région où, d'ailleurs, ainsi que je l'ai dit plus haut, persiste encore un dialecte aymará, le Kauki.

Actuellement, on ne doit plus classer dans la famille Aymará que les tribus suivantes, qui seules ont conservé leur langue

primitive : les Kol'a, terme générique qui désignait l'ensemble des tribus vivant autour du lac Titicaca ; les Lupaka, établis à l'Ouest du Titicaca jusqu'au Desaguadero au Sud, dans la province actuelle de Chucuito ; les Kol'agua, au Nord-Ouest d'Arequipa, sur le río Colca et aux sources du río Vitor ; les Pakase ou Pakaxe, qui occupaient toute la rive orientale du Titicaca et la région située au Sud de ce lac jusqu'à Callapa ; les Karanga ou Karanka, au Sud des précédents, entre le Desaguadero et le lac Coipasa ; les Tšarka, au Nord-Est du lac Poopó, où une province bolivienne porte leur nom (Charcas) ; les Kil'agua ou Kil'aka, qui doivent être la tribu la plus méridionale, car il existe, sur la rive sud du lac Poopó, une montagne qui porte leur nom (Quillacas), et enfin les Kauki, dont j'ai fixé l'habitat plus haut.

Si l'on voulait compléter cette liste avec les noms des tribus ayant parlé autrefois Aymará, mais ayant adopté maintenant le Kitšua, il faudrait y ajouter les *Kana, les *Kantši, les *Tšumbivilka, les *Aymará de la province des Aimaraes, les *Tšanka, les *Lipes, les *Tšitšas, que j'ai classés parmi les membres de la famille Kitšua (voir Famille Kitšua).

XI. Famille Bororó (**11** ; **40**).

La famille Bororó comprend :

a) les Bororó ou Coroados, qui occupent, dans le centre du Matto-Grosso, le haut Paraguay et ses affluents, le Jauru et le Cabaçal, le cours du São Lourenço jusqu'à son confluent avec le Cuyabá, atteignant au Nord le río dos Mortes, habitent aussi les deux rives de l'Araguaya jusqu'à 15° de latitude environ, et ont eu une colonie sur le río das Velhas, affluent sud du Paranahyba ;

b) les *Otukę, entre 17° et 18° de latitude et vers 59° de longitude ;

c) les *Kovareka, par 17° de latitude et 60° de longitude ;

d) les *Kuruminaka, par 16° de latitude et 60° de longitude ; et probablement aussi :

e) les *Korabeka, vers le 18ᵉ degré de latitude et entre les 60ᵉ et 61ᵉ degrés de longitude ;

f) les *Kuravę du río Tucabaca, affluent de l'Otuquis ;
h) les *Kurukaneka, voisins des Kuruminaka ;
i) les *Tapii, dont l'habitat se confond avec celui des Otukę.

XII. Famille *Diagit (Katamareño de Brinton, Kaltšaki de Chamberlain) (**6**, 12-32, 55-57).

Les *Diagit ou *Kaltšaki, dont la langue, le Kakan ou Kaka, encore vivante au XVIIe siècle, est inconnue, mais dont la belle civilisation a été étudiée avec le plus grand soin, occupaient toute la région montagneuse de l'Argentine, depuis le nevado d'Acay et la vallée de Lerma au Nord, probablement jusqu'à la province de Mendoza au Sud. Il faut sans doute rattacher aux Diagit les *Lule de la sierra d'Aconquija.

XIII. Famille Enimaga (**23**, 238-256 ; **31**).

C'est par erreur que Chamberlain (**9**, 239-240) réunit cette famille et la famille Maskoi. En effet, la langue parlée par les tribus des *Enimaga, des *Lengua, des *Guentuse et des Towothli est complètement différente de la langue parlée par les tribus de la famille Maskoi, encore qu'une de celles-ci soit désignée par les Espagnols sous le nom de Lengua.

Les *Enimaga, qui s'appelaient eux-mêmes Kotšabot ou Kotšaboth, vivaient, au moment de l'invasion espagnole, entre les ríos Bermejo et Pilcomayo, dans le Chaco central. Réduits à 150 guerriers, ils se retirèrent devant les Blancs et allèrent s'établir, à côté des Lengua de la famille Maskoi, aux sources du río Verde, affluent de droite du Paraguay. Les *Guentuse, qui leur étaient étroitement apparentés, les suivirent dans cette migration. Les *Lengua, qui s'appelaient eux-mêmes Juiadjé ou Oujadjé, formaient, au milieu du XIXe siècle, une très petite tribu, qui vivait, mélangée avec des Matšikui ou Maskoi et des Enimaga, un peu au Nord du Pilcomayo. C'est dans cette région, aux sources du Montelindo, que Grubb a retrouvé les seuls survivants de cette famille, les Towothli, Tóóthli, ou Toósle[1].

1. Les Toósle étaient classés jusqu'ici dans la famille Maskoi.

XIV. Famille *Esmeralda.

Les *Esmeralda occupaient autrefois tout le territoire compris entre le cours de l'Esmeraldas, le pays barbakóa et le Pacifique, et s'étendaient au Sud jusqu'au cap Pasado et peut-être plus bas encore le long de la côte.

XV. Famille Guahibo (33).

Les Guahibo (Goahivo, Goagivo, Uajibo, Guajiva, Guayba, Guayva) habitent les deux rives du Meta et la région comprise entre ce fleuve et le río Arauca d'une part, le río Vichada d'autre part ; à l'Est, ils atteignent la rive gauche de l'Orénoque.

Leurs principales tribus sont :

a) les Kuiva, Kuiba, Mella ou Ptamo, sur le bas Meta ;

b) les Tširikoa, sur le Lipa et l'Ele ;

c) les Katarro, sur le Yucabo ;

d) les Kuiloto, sur le río Cravo ;

e) les Tšuroye ou Bisanigua (Guaigua des anciens auteurs) et les Pamigua, dans le bassin du Guaviare, auxquels il faut rattacher probablement les Yamu du bas Ariari.

Anciennement, les Guahibo ont dû s'étendre jusqu'au Guayabero, dont le nom primitif « Papa-mene » semble dérivé de leur langue.

XVI. Famille Guarauno (77).

Les Guarauno (Uarao, Uarauno, Warrau) habitent le delta de l'Orénoque et le territoire adjacent à ce delta, ainsi qu'une partie du Nord-Ouest de la Guyane britannique.

XVII. Famille Guató (73).

Les Guató habitent, sur le haut Paraguay, la région des lacs Uberabá et Gaiba, ainsi que les bords de la lagune et de la rivière Caracara.

XVIII. Famille Guaykurú (28 *bis*).

La famille Guaykurú comprend un grand nombre de tribus

répandues sur les rives du Paraguay, du Paraná et de leurs affluents, et dans le Chaco.

a) Les *Mbayá-Guaykurú habitaient primitivement le Chaco boréal, atteignant au Sud Villa Hayes ; au milieu du XVIIe siècle, ils passèrent sur la rive gauche du Paraguay, où ils se transformèrent rapidement en une tribu de cavaliers redoutables, qui étendit son pouvoir sur tout le Paraguay septentrional, jusqu'au río Ypané au Sud (à peu près entre les parallèles 19° 28′ et 23° 36′, d'après Martius), poussant ses incursions jusqu'à Asunción au Sud, et Cuyabá au Nord. Leurs descendants actuels, les Kadiuéo, occupent un territoire limité à l'Ouest par le río Paraguay, au Sud par le río Branco, à l'Est et au Nord-Est par le río Miranda.

b) Les *Guatši vivaient, dans la seconde moitié du XIXe siècle, sur le río Miranda, au Nord des Kadiuéo, au milieu des Guaná, tribu arawak ; ils parlaient un dialecte très corrompu du Guaykurú, ou un dialecte étranger fortement mélangé d'éléments guaykurú.

c) Les Payaguá (appelés par certains auteurs Lengua) occupaient, au moment de l'invasion espagnole, un immense territoire le long du río Paraguay, dont ils étaient les pirates redoutés, poussant leurs incursions vers le Nord jusqu'à Cuyabá. Ils étaient alors divisés en une horde septentrionale, les Sarigué, par 21°5′ de latitude, dans la région occupée actuellement par les Kadiuéo, et une horde méridionale, les Magatš (nom dont les Espagnols firent Agaz, Agaces), par 25° 17′ de latitude, connus, à la fin du XVIIIe siècle, sous le nom de Siakuá ou Takunbú. En 1740, les Takunbú, en 1790, les Sarigué furent réunis à Asunción, où 40 à 50 de leurs descendants dégénérés vivent encore dans le quartier du port.

d) Les Toba, qui s'appelaient eux-mêmes Toko'it, Tokowit, et que les Espagnols désignaient sous le nom de Frentones ou Frontones, habitaient, à la fin du XVIIIe siècle, le Chaco central entre les ríos Pilcomayo et Bermejo, étendant leurs terribles incursions aussi bien dans le Chaco du Nord que dans le Chaco du Sud. Jusqu'à la fin du XVIIIe siècle, ils occupaient aussi le bassin du río San Francisco, aux sources du Bermejo, où ils furent remplacés à cette époque par des Matako. Actuellement, ils vivent en nombre

considérable sur les deux rives du Pilcomayo, d'où ils pénètrent profondément dans le Chaco boréal, où ils sont mélangés à des Matako (Guisnay, Tšorotí, Noktén, etc...) et à des Guaraní (Tapuí, Tapieté). Malgré son immense extension, la langue Toba est une.

Très proches parents des Toba sont les Pilagá, qui habitent la rive gauche du bas Pilcomayo, dans le Chaco boréal, et les *Aguilot, primitivement installés sur les rives du Bermejo, d'où ils émigrèrent, dans la seconde moitié du XVIIIe siècle, vers le Nord et se fondirent avec les précédents.

e) Les Mokoví (Mbokobí, Moskoví), dont le nom primitif semble être Mokouitt ou Mokovit, habitaient, à la fin du XVIIIe siècle, les deux rives du Bermejo, d'où ils pénétraient profondément dans le Chaco. Une partie d'entre eux émigra vers le Nord et se mêla aux Toba ; des restes insignifiants de la tribu vivent encore çà et là dans les parties reculées de la province de Santa Fe. Le Mokoví est un dialecte du Toba, altéré par des influences étrangères.

f) Les *Abipon (Callagá, Callage, Quiloaza) vivaient, vers le milieu du XVIIe siècle, sur la rive nord du Bermejo, d'où ils émigrèrent vers le Sud au début du XVIIIe siècle. Ils errèrent dès lors à travers l'immense Chaco austral depuis le Bermejo jusque dans les provinces de Santa Fe et de Paraná et jusqu'aux environs de Córdoba ; certains d'entre eux vinrent même s'installer dans la province de Corrientes en 1770. En 1858, il y avait encore une petite troupe abipon installée à un jour de marche de Santa Fe sur le chemin de Córdoba, et il se peut que quelques misérables restes de cette puissante tribu subsistent dans les plaines encore mal connues qui s'étendent entre Santa Fe et Santiago del Estero.

A la famille Guaykurú, je rattache, avec doute, les *Kerandí, qui vivaient le long de la rive droite du Paraná et du río de la Plata, notamment sur l'emplacement actuel de la ville de Santa Fe, atteignant à l'Est la sierra de Córdoba, et poussant leurs incursions au Sud-Est jusqu'à Buenos Aires et au río Salado (environ par 36° de latitude).

XIX. Famille *-Het (38).

Dans cette famille, Lehmann-Nitsche range les *Tšetšehet, tribu nomade qui vaguait entre la ville de Buenos Aires et la côte méridionale de l'actuelle province de Buenos Aires dans les sierras de Balcaree et de Ventana, atteignant au Sud le río Colorado et même le bas río Negro, et une fraction des *Diuihet (voir Famille Araukan).

XX. Famille Huari (50, 363-389).

Les Huari habitent, dans la cordillère des Paressis, aux sources du río Curumbiara, affluent de droite du Guaporé.

XXI. Famille Itonama (60).

Les Itonama habitent les rives du río de ce nom, depuis le lac Itonama presque jusqu'au confluent avec le río Machupo, entre les 13e et 14e degrés de latitude.

XXII. Famille Kahuapana (Mayna de Brinton).

La famille Kahuapana occupe un important territoire, sur les deux rives du haut Amazone, entre le Chambira et le Huallaga en aval, le Pastaza et le Potro en amont. On connaît trois dialectes de cette famille, le Xébero, le Mayna et le Kahuapana.

XXIII. Famille Kaliána (33).

Les Kaliána vivent sur le haut Parauá, affluent de gauche du Caróni.

XXIV. Famille *Kañari.

Les *Kañari occupaient toute la partie de la vallée interandine de l'actuelle République de l'Équateur, correspondant aux provinces de Cañar et de l'Azuay, et atteignaient presque à l'Ouest le littoral du Pacifique, entre les villes de Machalá et de Guayaquil.

XXV. Famille Kanitšana (10).

Les Kanitšana habitaient les rives du Mamoré, près des sources

du río Machupo, et le cours de ce dernier entre les 13ᵉ et 14ᵉ degrés de latitude.

XXVI. Famille Karadžá (Karayá) (34 *bis* ; 34 *ter*).

Les Karadžá vivent sur l'Araguaya entre les 15ᵉ et 6ᵉ degrés de latitude ; ils sont divisés en trois hordes : les Šambioá, au niveau des rapides de l'Araguaya ; les Žawažé, Žavahé ou Šavayé, dans les deux tiers nord de l'île Bananal ; les Karayá, sur la rive gauche, au Sud et à l'Ouest de l'île Bananal, d'où ils atteignent le Xingú entre les parallèles 5° et 8° 30′ environ.

XXVII. Famille Karib (2 ; 5 ; 18 ; 33 ; 44 ; 52 ; 67 ; 76 *bis* ; 77).

Le centre de dispersion de la famille Karib paraît avoir été la région comprise entre le haut Xingú et le Tapajoz, entre 10° et 12° de latitude. De là, leurs diverses tribus se sont répandues en éventail, vers le Nord, le Nord-Ouest et le Nord-Est, sur la moitié septentrionale du continent et une partie des Antilles.

Les représentants les plus méridionaux de la famille sont les Bakaïrí, qui habitent d'une part aux sources du Xingú sur les ríos Kulisehú (entre 12° 56′ et 13° 18′) et Tamitatoata ou Batový (entre 12° 30′ et 13° 15′), d'autre part aux sources du Tapajoz, sur les ríos San Manuel (vers 11°), Paranatinga et Arinos (le long duquel ils descendent jusqu'au 12ᵉ degré environ) ; les Kayabí, sur le Paranatinga et son affluent, le río Verde ; les Nahukwá (Nahuquá), sur la rive droite du Kulisehú (au niveau du parallèle 12° 50′) et sur son affluent, le Kuluëne, vers 12° 20′ ; les Aruma ou Yaruma, sur le Suyámissú, affluent de droite du haut Xingú.

Les Arára (Ajujuré) vivent sur les deux rives du Xingú entre 3° 35′ et 3° 45′, atteignant presque à l'Est le Tocantins, au Nord le Curuá, et à l'Ouest le Tapajoz, aux sources du Mauhé entre le Tapajoz et le Madeira, enfin entre ce dernier fleuve et le Purús, au Sud du lac Autaz ; ils sont apparentés aux Yuma, qui ont été signalés dans les forêts du Jacaré et sur l'Ituxi, affluents de droite du Purús. Les Apiaká (Apingui) habitent actuellement sur la rive gauche du bas Tocantins, à Praia grande dos Arroios et non loin

de la cataracte de Guariba. Probablement identiques à la tribu de même nom qui vague au Sud des monts Trocara, entre le Tocantins et le Xingú, vers 6°-7° de latitude, ils se souviennent d'avoir quitté le bassin du Xingú au milieu du XIXe siècle.

Trois autres tribus Karib vivent encore au Sud de l'Amazone : ce sont les Palmella, sur la rive droite du Guaporé entre ses affluents, les ríos Mequens et Blanco ou S. Simão ; les Parirí, sur le haut Pacajá, affluent méridional de l'Amazone (État du Pará), et les Pimenteira, autrefois entre les sources du Piauhy et du Gorguea, ultérieurement à Querebrobó sur le río San Francisco.

Au Nord de l'Amazone, les tribus Karib forment un bloc compact, qui occupe presque toute la rive gauche de ce fleuve jusqu'au río Negro en amont, une grande partie du bassin de l'Orénoque, l'Est du Vénézuela et les Guyanes.

Sur la rive gauche de l'Amazone, entre son embouchure et le río Negro, on rencontre les Apalai (Aparai) sur le Jary, le Parú et le Curuá, sans doute identiques aux *Arakwayú (Araquajú) signalés sur le Parú par les anciens auteurs [1] ; les Pianokotó (Pianoghottó), aux sources du Trombetas et du Jamundá ; les Pauxi, sur la rive droite du Cuminá, affluent du Trombetas ; les Wayewé, sur le río Mapuere, affluent du Trombetas, d'où ils s'étendent jusqu'aux monts Curicuri, sans doute identiques aux Woyawai (Voyavay) [2] des sources de l'Essequibo ; les *Bonari, sur le Uatumá.

Dans le bassin du río Negro, on rencontre les tribus suivantes : les Yauaperý (Uaimiri, Uaimeri), proches parents linguistiquement des Bonari, sur le río du même nom, affluent de gauche du río Negro ; les Krišaná (Kirisaman), aux sources du Yauaperý ; les Makuší (Makusi), sur le haut río Branco depuis l'Uraricuera et ses affluents septentrionaux jusqu'au Rupununi, origine occidentale de l'Essequibo ; les Taulipáng, près du Roroima, d'où ils s'étendent au Sud et au Sud-Ouest sur le haut Cuquenam et le haut Majarý et jusqu'à l'extrémité orientale de l'île Maracá ; les

1. L'Arakwayú renferme de nombreux mots tupi, ce qui l'a fait classer souvent dans la famille Tupi-Guarani.

2. Le Woyawai est un mélange de Karib et d'Arawak.

Seregóng, sur le haut Cutinho, auxquels sont étroitement apparentés les Ingarikó, établis au Nord du Roroima ; les Purukotó, (Ipurukotó, Porokoto, Purigoto), les Wayumará (Waiyamara, Uayamara, Vayamara) et les Sapará, qui n'existent plus en tant que tribus isolées, autrefois sur l'Uraricuera, branche occidentale du río Branco, d'où les Purukotó pénétraient sur le haut Paraguá ; les *Paravilhana (Parivilhana, Parauana, Paraguano, Šilikuna), autrefois sur le Caratirimany, affluent de droite du río Branco, postérieurement sur l'Uraricuera, puis sur le Mahú et aux sources du Tacutú, près des monts de la Lune, où le dernier représentant de la tribu est mort en 1914.

Dans le bassin de l'Orénoque vivent : les Makiritáre (Majongkong, Maiongcong, Majuyonco, Uayungomo), sur le haut Caurá-Merevari, aux sources de l'Auarý, sur le haut Ventuarí et les affluents de droite du haut Orénoque, le Cunucunúma, le Uapó (Iguapó), le Padámo, divisés en Yekuaná (Mayonggóng proprement dits), aux sources du Caurá, Ihuruána, aux sources du Ventuarí, Dekuána, sur les affluents de gauche du moyen et du bas Ventuarí, et Kunuaná, sur le Cunucunúma ; les Yabarána, actuellement réduits à 25 ou 30 individus sur la rive droite du cours moyen du Ventuarí, dont ils furent autrefois la tribu principale, qui parlent une langue très différente du Makiritáre, avec deux dialectes, le Kurašikiána en usage aux sources du Biehíta, affluent du Suapúre, et le Wökiáre aux sources du río Páro, affluent du Manapiári ; les Mapoyo, entre les ríos Paruaza et Suapúre, affluents de droite de l'Orénoque, qui sont les anciens *Kuakua ou *Nepoyo ; les Arinagoto, sur le Parágua entre 5° et 6° et sur le haut Caróni ; les Taparíto et les Panáre, entre le Caurá et le haut Cuchivero, les premiers sur le río Nicare, les seconds sur le río Mato, tous deux affluents de gauche du Caurá ; les Kariniako, sur l'Orénoque, non loin de l'embouchure du Caurá ; les *Tamanak, au Sud de l'Orénoque entre le Cuchivero et le Caróni ; les Arekuna (Jarekuna), au Nord des Makuší, aux sources du Caróni et de son affluent, le Parágua (où vit la tribu des Kamarakotó), sur le Mazaruni, affluent de gauche du bas Essequibo, surtout dans les environs du Roroima,

sur les affluents du haut Cuyuní, notamment sur le río Supamu, affluent du Yuruán.

La côte même du Venezuela était occupée, au moment de la découverte, par des tribus Karib. Les *Kumanagoto y vivaient depuis l'extrémité de la péninsule de Paria jusqu'au cap Codera et pénétraient à l'intérieur jusqu'aux plateaux de Barcelona ; ils comprenaient les *Tamanako (sans doute identiques aux Tamanak de l'Orénoque), aux sources du río de même nom, affluent de gauche de l'Unare, et du Manapire, affluent septentrional de l'Orénoque ; les *Tšaima ou *Uarapitše, dans les forêts du Guácharo et sur les ríos Guarapiche et Amana ; les *Tšakopata, entre le río Manzanares et la côte et entre les ríos Guere et Aragua ; les *Piritú, au Sud du petit port qui a conservé leur nom ; les *Palenke, sur la rive gauche de l'Unare au Nord du río Tamanaco ; les *Pariagoto, dans la presqu'île de Paria ; les *Kuneguara, entre Caripe et Maturín ; les *Guaikerí (Uaiquerí, Uiquire), dans la presqu'île d'Araya et l'île Marguerite. Les derniers survivants de cette importante tribu, appelés Karib, vivent actuellement dans la partie méridionale de l'État Anzoátegui et dans une partie de l'État Monagas, dans un territoire qui a pour limite à l'Ouest les ríos Unare et Suata, au Sud l'Orénoque, à l'Est le delta de ce fleuve occupé par les Guarauno ; les plus septentrionaux, qui vivent au Sud de Barcelona, sont très métissés. A ces tribus, je rattache avec doute les *Siparikot ou *Tšipa, qui vivaient, au Nord du Yaracuy, sur le río Aroa.

La Guyane est en grande pa tie peuplée de Karib.

Les Oyana (Ayana, Uayana), appelés au Brésil Urucuiana, en Guyane française Rukuyen, et par les tribus nègres de Surinam Alukuyana, vivent au Brésil sur le cours supérieur du Jary et du Parú et leurs affluents, en Guyane française sur le haut Lawa et ses affluents, en Guyane hollandaise sur le Paloumeu et le Tapanahoni. Sur le Paloumeu, ils sont mélangés à la tribu des Upurui, qui parle la même langue. Aux sources du Mana et du Sinnamarie, vivent les Taira (Guyane française). Les Yaho (Yao) habitent sur l'Ivaricopo (Guyane portugaise) et sur la rivière de Kaw. Les Trio, installés au XVIIIe siècle tout le long du Tapana-

honi, sont actuellement sur le haut Paloumeu, aux sources du Corentyn et sur le cours supérieur des affluents de l'Amazone qui prennent naissance dans la même région. Les Partamona vivent sur le Potaro, affluent de gauche de l'Essequibo. Les Uaika (Waika) habitent sur le Cuyuní et ses affluents (région du Yuruari) au Venezuela, sur le Barama en Guyane anglaise. Les Akawaí (Akawoio, Akauayo) occupent les rivières Pomeroon, Moruca, Cuyuní et Acarabisi en Guyane anglaise. Les Kaliña, enfin, appelés par la population créole des Guyanes Galibi, Karibi, Karaïb ou Caraïbes, étaient établis, à l'époque de la découverte, sur les cours moyen et inférieur des fleuves de la Guyane, depuis l'Oyapok jusqu'à l'Orénoque, vaste région, où ils ont survécu en quelques points, notamment à l'Ouest du bas Maroni et à l'Est du bas Corentyn, en Guyane hollandaise. Au moment de l'arrivée des Espagnols, ils occupaient aussi toutes les petites Antilles depuis Trinidad et Tobago jusqu'à Porto-Rico, dont ils commençaient à envahir la partie orientale, et faisaient parfois des expéditions jusque sur la côte est de Saint-Domingue [1]. En 1660, l'Angleterre et la France reléguèrent les 6000 Caraïbes survivants à la Dominique et à Saint-Vincent. En 1795-1796, ceux de cette dernière île, fortement mélangés de sang nègre (Black Caribs, Caraïbes noirs), furent déportés dans l'île de Ruatan et à Trujillo, sur la côte nord du Honduras, où ils prospérèrent. Actuellement, ils occupent la côte nord du Honduras, le port de Livingstone à l'embouchure du río Dulce au Guatémala, et quelques points de l'extrémité sud du littoral du Honduras britannique. Quant aux Caraïbes de la Dominique, il en reste 500 environ dans une réserve de la côte est.

Des fractions de la famille Karib existent encore, à l'état de groupes isolés, dans la partie occidentale de l'Amérique du Sud.

Les Motilon, auxquels on rattache les *Kirikire de la rive sud du lac Maracaibo, vivent dans la sierra de Perijá, sur le haut Catatumbo et ses affluents, le río de Oro et le Tarra ; autrefois, ils habitaient entre les ríos Catatumbo et Zulia et atteignaient les

1. Une de leurs tribus, les Cofachites, aurait même pénétré en Floride, chez les Apalatši.

environs des villes de Cucuta, Ocaña et Tamalameque. Leurs tribus principales sont les Yukuri, établis sur la pente occidentale de la sierra de Perijá, dans la région de Socomba et du haut Maracá ; les Tšake, divisés en Apón, Aponcito, Makoa (Makoita, Mokoita), dans le bassin du río Apón, à l'Ouest, au Nord-Ouest et au Nord du centre de Machiques, en Río Negro, Parirí et Yasa, sur le haut río Negro, et en Tukuko (Tukuku), sur le haut Tucuco, affluent du Santa Ana ; les Tšaparra ; les Irapeño.

Un autre groupe colombien est constitué par les Indiens de l'Opón et du Carare, affluents de droite du Magdalena, au Nord et au Nord-Ouest de Velez.

Le grand groupe des Umáua se trouve entre le haut Yapurá et le haut Caiarý, surtout sur le cours supérieur de l'Apaporís. Il comprend : les Hianákoto, sur le Macáya, affluent de l'Apaporís, les Tsahátsaha, sur le Cunyarý et son affluent, le Mesai, les Karixona (Kaŀihóna), sur la rive gauche du haut Yapurá, et les Guake (Huake), sur le río de los Engaños ou Yari, une partie du Caguán et de l'Orteguasa. Les Umáua correspondent sans doute aux *Guagua signalés par les anciens voyageurs entre le Caquetá et le Guayabero.

Dans la région de Jaen, c'est-à-dire sur les rives de l'Amazone au point où ce fleuve s'infléchit brusquement vers l'Est à la sortie de la cordillère, et sur le cours inférieur des affluents qu'il reçoit à ce niveau, le Chamaya, l'Utcubamba, le Chinchipe et le Tabaconas, vivait autrefois la tribu Karib des *Patagon.

Enfin, je rattache à la famille Karib une série de langues très différenciées, habituellement rangées dans une famille spéciale, la famille Peba ; ce sont le Peba proprement dit, le Yagua et le Yameo, parlés sur la rive gauche de l'Amazone entre 75° et 69° 30′ de longitude, sur le bas Tigre, le Nanay, le bas Napo et son affluent le Mazan, l'Apayacu, l'Ambiyacu et le Chichita.

Malgré les intéressants essais de L. Adam (**1**) et de de Goeje (**18**), la grammaire comparée des langues Karib est à écrire. Tout ce que l'on peut faire actuellement, c'est de les répartir en un certain nombre de groupes de dialectes affines :

1er groupe : Akawai, Arekuna, Kamarakotó, Makuši, Sapará,

Taulipáng, Seregóng, Ingarikó, Paravilhana, Ipurukotó, Krišaná ;

2e groupe : Trio, Hianákoto, Umáua, Guake, Karixona, Pianokotó, Wayumará, Makiritáre (Yekuaná, Ihuruána, Dekuána, Kunuaná) ;

3e groupe : Kumanagoto, Palenke, Guaikerí, Tamanak, Tšaima, Oyana, Upurui ;

4e groupe : Bakaïrí, Arára, Aruma, Yuma, Apiaká, Parirí.

5e groupe : Bonari, Yauaperý.

6e groupe : Peba, Yagua, Yameo.

Quant au Kaliña (avec ses dérivés le Caraïbe des îles et du Honduras), il paraît tenir à la fois des dialectes des groupes 2 et 3.

XXVIII. Famille Karirí.

La famille Karirí occupe un vaste territoire au Nord et à l'Ouest du río San Francisco dans les États de Bahia, Pernambuco, Piauhy et Ceará.

Elle comprend les Karirí proprement dits (Kirirí, Kayrirí), établis à l'intérieur des terres sur le río San Francisco, d'où leurs tribus se répandaient jusqu'aux ríos Curú et Acaracú (État de Ceará) à travers la sierra Borborema et les trois sierras des Cayriris ; et les Sabuyá, réunis dans les missions jésuites au Sud et l'Ouest de Bahia.

XXIX. Famille Katukina (64).

La famille Katukina occupe, d'une façon plus ou moins continue, un immense territoire situé au Sud de l'Amazone entre 72° 30′ et 62° 30′ de longitude, 4° et 9° de latitude, coupé diagonalement par le Juruá.

Elle comprend :

les Tukundiapa ou Tukano dyapá, sur le río das Piedras et le río Itecoahy, affluent du Javarý ;

les Parawa, sur la rive gauche du bas Gregorio, à Santo Amaro ;

les Bendiapa ou Beñ dyapá, sur la rive gauche du Juruá, en face de Bomjardim ;

les Tawari ou Kadekili dyapá, entre San Felipe sur le Juruá et les sources du Jutahý, où ils prennent le nom de Kayarára (Kairara) ou Wadyo paraniñ dyapá ;

les Kanamari, qui vivent d'une part sur la rive gauche du Juruá, entre 67° 45′ et 70° de longitude, d'où ils atteignent les sources du Jutahý et du Biá, d'autre part sur la rive droite, depuis le Tarauacá jusqu'aux sources du Pauiny, et, au Sud de ce fleuve, jusqu'au Purús ;

les Burue̜, sur le Jutahý et son affluent de droite, le Biá ;

les Katukina, répartis en deux groupes : un premier groupe, appelé Pidá dyapá, établi sur le moyen Jutahý, et en particulier sur ses deux affluents le Mutum et le Biá, dont une fraction installée sur l'Igarapé preto, affluent de droite du Jandiatuba, porte le nom de Kutía dyapá ; un second groupe, qui s'étend de la rive droite du Tarauacá à la rive gauche du Purús au Sud du Tapauá, et dont une fraction vivait aux sources du Teffé ;

les Katawiši, qui s'appellent eux-mêmes Hewadie, dont on trouve des tribus entre le Purús et le Madeira entre 7° 30′ et 6° de latitude, entre le Purús et le Juruá au niveau du 5ᵉ parallèle, et enfin sur le Juruá même et sur sa rive gauche au Sud du Rio-sinho.

XXX. Famille Kayuvava (13).

Les Kayuvava habitent sur la rive occidentale du Mamoré, à une quinzaine de lieues au-dessus de son confluent avec le Guaporé, et sur les petits affluents des plaines de l'Ouest entre 12° et 13° de latitude sud.

XXXI. Famille Kitšua.

Le Kitšua ou Runa-simi est la seule langue de l'Amérique du Sud, qui ait joué, à l'époque précolombienne, le rôle d'une « langue de civilisation ». Propagée par les conquérants incasiques, elle se répandit peu à peu dans presque tout leur vaste empire, dont elle devint la langue officielle et où elle joua le rôle de langue de relation.

Cet empire, lors de l'arrivée des Espagnols, s'étendait depuis le río Angasmayo au Nord (à la limite des Républiques actuelles de Colombie et d'Équateur) jusqu'au río Maule (Chili) au Sud, englobant les territoires qui correspondent maintenant à la partie

andine et littorale de la République de l'Équateur (à l'exception de la province côtière d'Esmeraldas), à la région andine et littorale du Pérou, au Chili jusqu'au río Maule, à la haute Bolivie et à la région andine de la République Argentine. Toutefois, la domination incasique ne dura pas assez longtemps ou ne fut pas assez forte pour imposer le Kitšua à toutes les populations subjuguées. C'est ainsi que, lors de la conquête espagnole, des langues, différentes du Kitšua, étaient encore bien vivantes dans un grand nombre de provinces : le Pasto, le Kil'asinga, l'Atal'an, le Puruhá, le Kañari, le Palta en Équateur, le Sek, le Yunka, l'Atakama, le Tšango-Uru-Pukina, l'Aymará, au Pérou et en Bolivie, l'Araukan au Chili, le Kakan en pays diagit.

Sous la domination espagnole, l'expansion du Kitšua s'est accentuée d'une façon remarquable, grâce aux missionnaires qui, l'ayant adopté comme langue d'évangélisation, en imposèrent l'usage à leurs néophytes. Leur patient effort réalisa l'œuvre inachevée des conquérants incasiques ; c'est ainsi que le Kitšua supplanta définitivement le Puruhá, le Kañari, le Palta en Équateur, le Kakan en pays diagit et restreignit encore davantage le domaine de l'Aymará et de l'Uru-Pukina, cependant que, de son côté, l'Espagnol se substituait aux langues indigènes, surtout dans la région côtière. Bien plus, les missionnaires implantèrent le Kitšua dans des régions qui avaient échappé à la conquête des Incas ; c'est ainsi que, avec l'Évangile, la langue du Cuzco pénétra dans la province de Santiago del Estero, dans les bassins du haut Tuichi, affluent du Beni, du haut Amazone (dialecte Mayna), de l'Ucayali, du haut Napo, du haut Caquetá et du haut Putumayo (dialecte Ingano), et dans le Sud de l'État colombien de Tolima (dialecte Almaguero).

Il est indispensable de tenir compte de ces faits historiques, qui ont complètement transformé l'aspect linguistique de toute cette région de l'Amérique du Sud, au cours des quatre derniers siècles, et notablement élargi le domaine du Kitšua.

Les tribus parlant cette langue peuvent être divisées en cinq groupes géographiques : le groupe Inka, le groupe Tšintšasuyu, le groupe Kiteño, le groupe bolivien, le groupe argentin.

a) Le groupe Inka comprenait les Kana et les Kantši, installés dans la haute vallée de Vilcãnota, les premiers entre le col de Vilcanota au Sud et Cacha au Nord, les seconds entre Cacha et Urubamba ; les Inka proprement dits, installés entre l'Apurimac et le Paucartambo, dans un territoire que divise longitudinalement la vallée de Vilcanota, vallée qu'ils occupaient depuis Ollantaïtambo au Nord jusqu'à Quiquijana au Sud ; les Tšumbivilka, installés sur les ríos Santo Tomas et Velille, affluents de gauche du haut Apurimac ; les Aymará, sur le haut Pachachaca, affluent de gauche du haut Apurimac ; les Kitšua, primitivement établis entre l'Apurimac et le río Pampas, dans les vallées d'Abancay et d'Andahuaylas, puis refoulés vers le Sud par les Tšanka dans la région montagneuse où prennent leurs sources l'Apurimac, le Pachachaca et leurs affluents ; les Tšanka, tout d'abord installés aux environs d'Ayacucho et de Huanta avec comme limite orientale le río Pampas, qui, après avoir refoulé les Kitšua, s'étendirent iusqu'à l'Apurimac ; les Huanka, qui occupaient la vallée de Jauja et les rives du lac de Junín (Pumpu ou Chinchaycocha) jusqu'au sommet du cerro de Pasco au Nord ; les Rukana, qui vivaient au Sud-Ouest des Tšanka dans la cordillère occidentale et sur ses deux versants vers le parallèle 15°.

Un certain nombre de ces tribus parlaient primitivement l'Aymará et n'ont adopté le Kitšua qu'à une époque relativement récente : ce sont les Kana, les Kantši, les Tšumbivilka, les Aymará, les Tšanka et peut-être les Kitšua eux-mêmes (voir Famille Aymará).

b) Le groupe Tšintšasuyu occupait la double cordillère des Andes, avec la vallée du Marañón qu'elle borde à l'Est et à l'Ouest, depuis le cerro de Pasco au Sud, jusqu'au río Macará (frontière péruano-équatorienne) au Nord. Il comprenait les Huanuku, la peuplade la plus méridionale du groupe, dans la région de la ville actuelle de Huánuco ; les Kontšuku, dans la région de Huaraz, aussi bien dans les vallées du versant du Marañón que dans celles du versant du Pacifique ; les Huamatšuku, plus en aval dans la vallée du Marañón, où le village de Huamachuco conserve leur nom ; les Kasamarka, un peu en aval des précédents,

dans une région qui doit sans doute correspondre aux centres actuels de Cajamarquilla et de Cajamarca; les Tšatšapuya ou Tša-tša, dans la contrée montagneuse de la rive droite du Marañón, où se trouve maintenant la ville de Chachapoyas; les Huakra-tšukru, de chaque côté du défilé du Marañón; les Huankapampa, dans les montagnes près de Jaen; les Ayahuaka, sur le haut Quiros, affluent du río de la Chira, dans la région où se trouve la ville de Ayavaca. A ce groupe, se rattachent les Lamaño ou Lamista, installés autrefois aux environs de Lamas, sur le río Mayo, affluent de gauche du Huallaga, dans l'ancien diocèse de Truxillo.

c) Le groupe Kiteño constitué, à l'époque de la découverte, uniquement par les Kara de la région de Quito, qui, dès cette époque, quoique depuis peu d'années, avaient renoncé à leur idiome originel pour adopter le Kitšua, comprend actuellement toutes les peuplades de la région andine équatorienne.

d) Le groupe bolivien comprend les peuplades de parler kitšua des départements de Cochabamba (provinces de Sacaya, d'Ayopaya, de Tacapari, d'Arque, de Clisa, de Mizque), de Chuquisaca (provinces de Yamparais, de Tomina, de Sinti) et de Potosi (provinces de Potosi, de Chayanta, de Porco, de Chichas et de Lipes). Les Tšitšas et les Lipes parlaient primitivement aymará (voir Famille Aymará).

e) Le groupe argentin comprend toute la région andine de la République Argentine (territorio de los Andes, provinces de Jujuy (Omaguaka), de Salta, de Tucuman, de Catamarca, de La Rioja et de San Juan), et même, en dehors de la cordillère, la province de Santiago del Estero. La substitution du Kitšua au Diagit ou Kakan était déjà en partie réalisée au moment de la découverte; elle est devenue complète à la fin du XVIIe siècle, époque où le Kakan paraît avoir disparu. La substitution du Kitšua à l'Al'entiak, au Sanaviron, au Vilela-Tšulupí est postérieure à la conquête.

Les principaux dialectes du Kitšua sont : le Kiteño, parlé en Équateur; le Lamaño ou Lamista; le Tšintšasuyu ou Tšintšaya, avec les sous-dialectes de Huari (département d'Ancachs), de Huánuco, peu différent du précédent, et de Cajamarca, qui

n'est plus en usage que dans la Pampa de los Baños del Inca, près de Cajamarca ; le dialecte de la province de Huancayo (département de Junín) ; le dialecte d'Ayacucho ; le Kuskeño, parlé dans la région du Cuzco ; le dialecte bolivien (départements de Cochabamba et Chuquisaca) ; le dialecte argentin ou Tukumano, appelé par les indigènes Kuzko.

On possède également des documents sur les parlers kitšua, en usage dans des régions qui n'appartenaient pas à l'empire incasique, par conséquent implantés grâce à l'action des Blancs : dialecte Ingano, dialecte Almaguero, dialecte du Napo, dialecte Mayna, dialecte de l'Ucayali, dialecte du Tuichi, dialecte de Santiago del Estero.

XXXII. Famille Kotše (Mokóa de Brinton) (**8** *bis* ; **67**).

La langue Kotše n'est plus parlée actuellement que dans un petit village indien de la Colombie orientale, Sebondoy, par les derniers représentants de la tribu des *Mokóa. Ces Mokóa semblent identiques aux anciennes tribus des *Patoko et des *Kil'asinga, qui habitaient toute la cordillère orientale des Andes depuis, et y compris, la lagune La Cocha au Sud, jusqu'au páramo de las Papas au Nord, et occupaient le haut Putumayo et le haut Caquetá à l'Est.

XXXIII. Famille *Kófane (**81**,61-62).

Les *Kòfane vivaient à l'Est et près de la montagne Cayambe, aux sources de l'Aguarico, entre ce fleuve et le río Azuela et sur les rives des ríos Cofanes, Sardinas, Duino et Payamino. Une de leurs tribus, établie sur les bords de la lagune de Cuyabeno, portait le nom de *Mako.

XXXIV. Famille Leko.

Les Leko occupaient, sur la rive gauche du Beni, le bassin du Guanay ou Kaka et de ses affluents, le Tipuani, le Mapiri, le Turiapo, le Yuyo, etc... ; leur langue s'appelle le Lapalapa.

XXXV. Famille Mǻku (**33**).

Les Máku habitent sur le río Auarý, affluent de gauche du haut Uraricuéra (une des branches du río Branco).

Il ne faut confondre ces Indiens ni avec les Makú du río Negro et du Yapurá, qui sont des Puináve, ni avec les Máku du Ventuarí, qui sont des Sáliba, ni avec les Mako de la lagune de Cuyabeno, qui sont des Kófane.

XXXVI. Famille Maskoi (Machicui ou Muscovi de Ehrenreich ; Enimaga de Chamberlain) (**9** *ter*; **19**, 318-321 ; **31**).

La famille Maskoi occupe une large bande de territoire à l'Ouest du Paraguay, depuis le confluent de ce fleuve avec le Pilcomayo au Sud, jusqu'au 20e degré de latitude au Nord. Bordant la rive droite du Paraguay jusqu'au parallèle 22°, elle s'écarte ensuite de ce fleuve dans la direction du Nord-Ouest.

Cette famille comprend : les Maskoi, près du confluent du Paraguay et du Pilcomayo, depuis ce fleuve jusqu'à la petite rivière Montelindo, dans les environs de Villa Hayes ; les Lengua ou Gekoinlahaák, entre le Montelindo et le parallèle 22° 45′ ; les Angaité (Enslét), sur la rive droite du Paraguay, en face de l'ancienne mission San Salvador ; les Sanapaná, sur la rive droite du río Salado ; les Sapukí et les Guaná [1], au Nord-Ouest des précédents, jusqu'au parallèle 20°.

XXXVII. Famille Mašubi.

Les Mašubi vivent à l'Est du Guaporé, sur le cours moyen ou supérieur du río Mequens. Leur langue ne nous est connue que par un court vocabulaire, encore inédit, recueilli par le colonel Fawcett, qu'a bien voulu me communiquer E. Nordenskiöld.

XXXVIII. Famille Matako-Mataguayo (**22**; **23** ; **48**, 14-136).

Cette famille forme un groupe compact dans le moyen Chaco, à l'Ouest des Toba (famille Guaykurú).

Elle comprend :

les Mataguayo proprement dits, installés aux sources du Bermejo au pied de la cordillère orientale des Andes ;

1. Il faut soigneusement distinguer ces Indiens de leurs homonymes de la famille Arawak.

les Matako, à l'Est des précédents, sur la rive droite du Bermejo ;

les Vexoz, qui s'appellent eux-mêmes Aiyo, installés sur la rive gauche du Bermejo entre ce fleuve et le Piquirenda ;

les Huešuo, les Abutšeta, les Pesatupe, les Imaka, à l'Est des précédents, entre le Bermejo et l'Itiyuro au Nord ;

les Noktén, au pied de la cordillère de Pirapo, entre le Pilcomayo, le Piquirenda et l'Itiyuro ;

les Guisnay, qui sont la continuation orientale des Noktén ;

les Tšorotí (Tsoloti, Soloti, Solote, Tšorote), qui s'appellent eux-mêmes Yófuaha, installés en face des Noktén et des Guisnay sur la rive gauche du Pilcomayo, entre 21° 30′ et 22° 30′ de latitude, depuis le fort Guachalla en aval jusqu'à San Francisco ou Villa Montes en amont ;

les Ašluslay, Suhín ou Súxen [1], qui vivent sur le Pilcomayo et sur sa rive septentrionale, depuis le pays des Noktén et des Tšorotí jusqu'au pays occupé par la famille Maskoi ;

les Malbalá, qui forment une petite enclave en territoire Vilela-Tšulupí, sur la rive droite du Bermejo.

XXXIX. Famille Mobima (**12**).

Les Mobima (Movima) habitaient les plaines de l'Ouest du Mamoré, sur les rives du río Yacumá et de son affluent, le Rapulo.

XL. Famille Moseten (**4** ; **50**, 181-267).

La famille Moseten occupe la région montagneuse à l'Est du Beni, entre 15° et 16° de latitude et 65° 30′ et 67° 30′ de longitude. On y classe les Moseten proprement dits (Maniki, Magdalenos, Tšumano, Tšomane), aux sources du Maniqui et du Securé, affluents du Mamoré, et du Quiquiré, affluent du Beni, et les Tšimáne, au Nord des précédents, sur le haut Rapulo.

XLI. Famille Múra (**46** *ter* ; **64**).

Les Múra habitaient le cours inférieur du Purús, la rive gauche

1. Les Súxen étaient classés jusqu'ici dans la famille Maskoi. Leur identité avec les Suhín ne me paraît pas douteuse. Par contre, c'est par erreur que Hunt identifie les Suhín et les Tšulupí (**23**, 257).

de l'Amazone, entre le bas río Negro et le Yapurá, et divers points de la rive droite du même fleuve, notamment entre le Teffé et le Madeira, au niveau du lac Mamia et à l'embouchure du Jutahý. Actuellement, on ne trouve plus de représentants de cette tribu que sur le río Autaz (Buruaray), sur le Manicoré (Múra-Bohurá), sur le Maicy, affluent du Marmellos (Múra-Pirahá), et dans le bassin du río Caštanha ou Roosevelt (Matanauý), d'où quelques individus sont venus s'établir sur le bas Marmellos.

XLII. Famille Nambikuára (**69**, 151-152).

Les Nambikuára vivent sur le haut Juruena, affluent du Tapajoz, sur le haut río Roosevelt, affluent du Madeira, et sur le haut Guaporé. On y distingue :

a) un groupe du Sud-Est, constitué par les Kôkôzú ou Kôkôsú, sur les ríos Juruena, Juina, Papagaio et Camararé ;

b) un groupe du Nord-Est, constitué par les Anunzê du río 12 de Outobro, qui s'étendent probablement jusqu'au confluent de l'Arinos et du Juruena, sur la rive gauche de ce fleuve ;

c) un groupe du Sud-Ouest, constitué par les Uaintasú ou Uáindzê ou Kabiši, établis entre le río Camararé, affluent du haut Tapajoz, et le Guaporé ;

d) un groupe du Nord-Ouest, constitué par les Tagnanís des sources du río Roosevelt, divisés eux-mêmes en Tauitês, Salumás, Tarutês et Taschuitês.

XLIII. Famille Otomak.

Les Otomak vivent dans le Sud-Ouest du Venezuela, entre l'Orénoque, le Meta, le haut Arauca et le Sinaruco. L'Otomak n'avait, d'après les missionnaires, qu'un seul dialecte, le Tarapita.

XLIV. Famille Pano (**17** *bis*, 88-95, 110-113; **21**; **64**).

La famille Pano est divisée géographiquement en trois groupes : le premier groupe, le plus important, occupe la rive sud de l'Amazone, depuis le Jutahý à l'Est jusqu'au Huallaga à l'Ouest, le bassin entier de son affluent de droite, le Javarý, les deux rives de l'Ucayali depuis son embouchure jusqu'au parallèle 10°, et la

rive droite de son affluent, l'Urubamba, tout le bassin du haut Juruá et les sources mêmes du Purús ; le 2e groupe occupe le bassin de l'Inambari er de ses affluents, le Marcapata, le Yaguarmayo et le Chaspa ; le 3e groupe est installé sur les rives du Mamoré, du Beni et du Madre de Dios entre 9° 15′ et 12° 30′ de latitude, 64° 45′ et 67° 30′ de longitude.

Au premier groupe appartiennent les Kulino (Kurina), établis sur la rive méridionale de l'Amazone entre le Jutahý et le Javarý ; les Mayoruna (Maxuruna, Pelados), qui occupent le haut Jandiatuba, la plus grande partie du bassin du Javarý, atteignant au Nord l'Amazone, à l'Ouest le bas Ucayali, auxquels on rattache les Marubo (Maruba, Moruba, Marova), qui vivent sur le río de même nom et le río Cochiquina, et les Tširabo, dont l'habitat est limité par le Tahuaya, le Javarý chico, le Charapa et l'Amazone ; les *Itukale ou *Urarina, la seule tribu pano installée au Nord de l'Amazone, sur les rives du Chambira ; les Tšamikuro, sur le río de même nom, affluent de droite de l'Amazone, en amont de l'Ucayali, qui atteignent le Huallaga à l'Est ; les Kapanahua, qu'on trouve d'une part aux sources du Javarý, du Tapiche et du Blanco, sur le Huanachá et l'Alacrán (où ils portaient le nom de Buskipani), d'autre part dans le massif montagneux où prennent leurs sources le Tejo, le Gregorio, le Liberdade, le Tarauacá et le Breo ; les Katukina, répartis en deux groupes, l'un sur le Jaquiraná, aux sources du Javarý, l'autre sur la rive gauche du Gregorio (haut Juruá) et sur le río Katukina (haut Tarauacá) ; les Remo, dans les montagnes entre le Juruá et l'Ucayali, entre 8° 30′ et 7° de latitude, auxquels appartiennent les Sakuya des sources du Tamaya ; les Sensi (Sensivo), sur le Huanachá et le Chanuya ; les Šetibo (Setebo, Šitebo, Xitipo) ou Pano, sur le cours inférieur de l'Ucayali jusqu'au parallèle 7° au Sud ; les *Maparina, aux sources du Cuxiabatay ; les Sipibo (Tšipeo, Šipeo, Tšepeo, Tšipibo, Sépibo, Šipebo, Xipibo), sur l'Ucayali entre 7° et 8° 30′ de latitude et sur son affluent de gauche, le Pisqui (où ils portent le nom de Sinabu), auxquels se rattachent sans doute les *Kal'iseka (Kaliseka) de la rive gauche du Pachitea, et les Mananagua du bassin du Callaria ; les Kašibo (Katšibo, Kahibo), qui vivent

sur le Pachitea, le Pisqui et l'Aguaita, affluents de gauche de l'Ucayali, dont les sous-tribus sont les Buninahua, les Karapatšo et les Putšanahua ; les Kunibo, sur l'Ucayali, entre 8° 30′ et 10° de latitude ; les Pitsobu et les Saboibo, sur le Tahuania, affluent de droite de l' Ucayali ; les Amahuaka (Maspo, Impetineri), qui occupent les montagnes entre le bassin de l'Ucayali et de son affluent, l'Urubamba, et les bassins du Purús et du Juruá, entre 9° 15′ et 11° 30′ de latitude, ainsi que les rivières qui en descendent à l'Est et à l'Ouest, et dont certaines tribus ont pénétré dans le haut Juruá et dans le haut Purús ; les Yaminawa, sur le haut Envirá et aux sources du Tarauacá jusqu'au haut Juruá à l'Ouest ; les Kašinahua, sur la rive gauche de l'Envirá, le Tarauacá, le haut Gregorio et le haut Liberdade ; les Arara, Ararapina et Ararawa (sans doute identiques aux Saninawa), sur le haut Liberdade et l'Humayta ; les Nukuini, sur le haut Mõa ; les Kontanawa, sur le haut Tarauacá et l'Humayta ; le Kuyanawa, entre le Mõa et le Paraná dos Mouras ; les Marinawa, sur le Furnaya (haut Envirá) ; les Tušinawa, sur l'Humayta et le Furnaya ; les Pakanawa, aux sources de l'Envirá ; les Šipinaua, entre le haut Liberdade et le haut Valparaiso et sur l'Amoaca et le Grajahu, affluents de droite du haut Juruá ; les Yauavo (Jawabu), sur l'Acuria, affluent de droite du haut Juruá ; les Yura, sur le Piqueyacu et le Torolluc, aux sources du Juruá ; les Kanamari ou Kanawari, sur le haut Purús et son affluent de gauche, le Curumahá.

Dans le deuxième groupe, on classe les Arasaire ou Arasa [1], sur les bords du Marcapata, affluent de gauche de l'Inambari ; les Yamiaka, sur le Yaguarmayo, affluent de droite du même fleuve ; les Atsahuaka [1], sur le Carama, affluent de gauche du Tambopata, et le Chaspa, affluent de droite de l'Inambari ; les Araua, petite tribu isolée sur le Chive, affluent de gauche du Madre de Dios.

Rentrent dans le troisième groupe : les Pakaguara, installés sur les deux rives du Beni, du bas Madre de Dios, du Mamoré, du Madeira et de l'Abuná, qui sont divisés en plusieurs tribus : les Kapuibo dans le bassin du río Biata, les Tšakobo entre le

1. Ces deux peuplades sont bilingues et emploient aussi le Takana.

Mamoré et le lac Rogoaguado, les Sinabo aux environs des premiers rapides du Mamoré, et les Karipuna, établis aux environs des chutes du Madeira, qui comprennent eux-mêmes les Yakariá ou Yakaré-Tapuüya, sur l'Abuná, et les Šenábu, en amont de la cascade de Pao grande (probablement identiques aux Sinabo cités plus haut).

Je rappelle que les Takana, que je classe dans la famille Arawak, parlent une langue dont la grammaire offre de grandes ressemblances avec la grammaire Pano (voir Famille Arawak).

XLV. Famille Pueltše (-Künnü de Lehmann-Nitsche) (**38**).

Les Pueltše, appelés par Falkner Tehuelkünnü, Tehuelhet et Tehueltše, par Camaño Tueltše (branche australe), par d'Orbigny Pueltše, par Hale Pueltše, Pampa, Tehuilitše, par Cox Tehueltše du Nord ou Pampa, par Musters Pampa, par Moreno Gennaken, Pampa véritables ou Tehueltše du Nord, et dans les anciens documents Tueltše, Toeltše, habitaient dans l'Argentine centrale et méridionale depuis les derniers contreforts des Andes au Sud du Limay jusqu'à l'Atlantique, entre le río Negro et le 42e parallèle, atteignant même au Nord le río Colorado et au Sud le haut Chubut ; ils vivent maintenant surtout entre les ríos Colorado et Negro. Leur langue comporte deux dialectes, le dialecte oriental parlé près de Carmen de Patagones, le dialecte occidental en usage dans les campos de Maquinchao au Sud du río Limay.

XLVI. Famille Puinâve (**63**).

La famille Puinâve comprend :

a) les Puinâve (Puinabe, Puinavis, Uaipunabis, Guaipunavos, Uaipís), qui habitent le bassin de l'Inírida ;

b) les Makú, qui nomadisent entre le río Negro et le Yapurá entre 69°30′ et 61°45′ de longitude.

XLVII. Famille *Puruhá (**81**, 23-29).

Les *Puruhá habitaient la province actuelle du Chimborazo (Équateur).

XLVIII. Famille Sáliba (59).

La famille Sáliba comprend :

a) les Sáliba (Sáliva, Sálliba, Sáliua), qui occupaient originellement le territoire compris entre le Vichada, le Guaviare et l'Orénoque, et comme habitat secondaire, le río Muco, le haut Meta et le haut Vichada ;

b) les Piaróa, dont le centre principal se trouve sur le río Sipápo et la rive droite de l'Orénoque, dans les environs des rapides Átures et Maipúres ;

c) les Máku, installés dans les savanes de la rive droite du moyen et du bas Ventuarí, et aussi sur le cours supérieur des affluents de droite de l'Orénoque, depuis le confluent du Ventuarí jusqu'à quelques jours en aval de l'embouchure du Cunucunúma.

XLIX. Famille Samuku (48, 298-304).

La famille Samuku occupe un vaste territoire du Chaco septentrional, encore très mal connu, qui s'étend depuis le Paraguay à l'Est jusqu'au río Parapití à l'Ouest, depuis le río Otuquis au Nord jusqu'au parallèle 21° au Sud. Elle comprend : les Tšamakoko, la tribu la plus orientale (dont les *Kaipotorade ou *Kaipotade, les *Tunatšo et les *Imono, mentionnés par les anciens auteurs, ne sont sans doute que des fractions), qui vivent sur la rive droite du Paraguay depuis un point situé à peu près à égale distance entre Albuquerque et Coimbra jusqu'à un peu en aval de Ft. Olympo, les représentants les plus méridionaux de la tribu, restés à l'état sauvage, portant le nom de Tumanahá (Timinaha, Tumaná) ; les Guarañoka ou Guaramoka, à l'Est des Salines de San José et de Santiago ; les Morotoko, au Sud de San José et de San Juan, vers le parallèle 19°, divisés en Koroïno, Karera, Tomoeno ou Tameono, Kukurave ou Kukutade, Panono, Ororebate ; les Poturero, sur le río Otuquis et son affluent, le río San Rafael, auxquels il faut sans doute rattacher les *Ninakiguila de la rive droite du bas Otuquis ; les Samuku (Zamuko), qui vivaïent entre 20° et 21° de latitude et vers le 61ᵉ degré de lon-

gitude, avec une sous-tribu, les Zatieno (Satieno, Zatineo) ; les Ugaraño (Ugarono), grande tribu à l'Ouest des Samuku et au Sud des Tsirakuá, et enfin les Tsírakuá, que Nordenskiöld a rencontrés, divisés en deux groupes, entre les 19e et 20e degrés de latitude, à l'Est et à l'Ouest du Parapití, séparés l'un de l'autre par une tribu arawak guaranisée, les Tšané.

L. Famille *Sanaviron (**6**, 37-41).

Les *Sanaviron vivaient entre Córdoba et Santiago del Estero, c'est-à-dire à l'Est et au Sud des Salinas grandes de Córdoba. Il est probable que le *Kometšingon, parlé dans la sierra de Córdoba, était un dialecte du Sanaviron.

LI. Famille Šavanté (Eošavanté de von Ihering) (**26**).

Les Šavanté habitent, dans l'État brésilien de São Paulo, le cours inférieur des ríos Tieté et Paranapanéma, vers le parallèle 20°. Ils s'appellent eux-mêmes Oti et sont complètement différents linguistiquement des Šavánte Opaié de l'État de Goyaz, qui appartiennent à la famille Žẹ.

LII. Famille Sek (**8**).

La famille Sek, confondue à tort avec la famille Yunka, occupe comme celle-ci les vallées chaudes de la côte péruvienne ; son domaine s'étend entre 5° et 6° 30′ de latitude méridionale. Elle comprend les *Kolan sur le río de la Chira, au Nord de Payta, les Katakao, sur le cours supérieur du río Piura, qui, d'après Means, parlent encore leur langue primitive, et les *Setšura, sur le cours inférieur du même fleuve.

LIII. Famille Širianá (**33**).

Cette famille comprend :

a) les Širianá ou Širišána, qui vivent répartis en petits groupes, à l'intérieur des terres, sur les affluents des deux rives de l'Uraricuéra, sur le versant méridional des monts Marutaní aux sources de l'Uraricapará, et sur les deux versants de la grande cordillère de Parima, qui fait le partage des eaux entre l'Uraricuéra et l'Orénoque ;

b) les Waíka, qui vivent sur le haut Ocámo, affluent de droite du haut Orénoque.

Širianá et Waíka descendent des *Guaharibo, tribu autrefois très redoutée des sources de l'Orénoque.

LIV. Famille Timote (Muku de Salas).

La famille Timote comprend le groupe Timote et le groupe Kuika.

a) Les multiples tribus du groupe Timote occupaient l'actuel État de Mérida, c'est-à-dire le bassin du Chama, les deux cordillères qui le limitent au Nord et au Sud, et les cours supérieurs des rivières qui y prennent naissance et se rendent au lac de Maracaibo d'une part, à l'Apure d'autre part. A l'Ouest, le groupe englobait les sources de certains affluents de droite du Tachira. Les dialectes attestés sont le Timote, le Mirripú ou Maripú, le Mukutší ou Mokotší, le Migurí, l'Eskaguey et le Tiguiñó.

b) Le groupe Kuika occupait la plus grande partie de l'État actuel de Trujillo et comprenait les Kuika proprement dits, les Tostó, les Eskuke et les Xaxó.

LV. Famille Trumaí.

Les Trumaí habitent sur la rive gauche du bas Kulisehú, affluent du haut Xingú, par 12° 5′ environ de latitude.

LVI. Famille Tšapakura (46 *ter*; 50, 399-433).

Cette famille comprend :

a) les Tšapakura ou Huatši, installés sur les rives du río Blanco ou Baures, à peu près par 15° de latitude et 62° de longitude ;

b) les Kitemoka, originaires des mêmes régions que les précédents, dont une des tribus était les Napeka ;

c) les Pawumwa ou Huanyam, installés dans le bassin du río São Miguel, affluent de droite du Guaporé, et sur la rive droite de ce dernier fleuve en aval et à l'intérieur des terres ;

d) les Iten, entre les ríos Guaporé et Mamoré, près de leur confluent ;

e) les Turá du río Marmellos, les Arára de l'embouchure du

río Preto, affluent du Madeira, les Urupá et les Jarú des ríos de même nom, affluents de gauche du haut Machado, et les Arikem du haut Jamary, affluent du Madeira.

A cette famille, il faut vraisemblablement rattacher aussi : les *Mure, les *Rokorona, les *Rokotona, les *Orokotona, les *Rotoróño, les *Okoróno et les *Herisobokono ou *Herisibokóno, répartis en deux groupes, l'un sur la rive droite du Baures depuis son confluent avec le Guaporé au Nord jusqu'à Concepción de Baures au Sud, groupe qui correspond aux *Sansimoniano de certains auteurs, l'autre sur la rive gauche du Mamoré, aux sources de ses affluents, le Rapúlo ou Maniqui et le Tijamuchi.

LVII. Famille *Tšarrúa (55 ; 75).

La famille *Tšarrúa occupait la région comprise entre le Paraná et la côte, l'embouchure du río de la Plata et la laguna dos Patos. On y classe les *Tšarrúa proprement dits, les *Güenoa, les *Yaro, les *Bohane, les *Minuán, les *Tšaná-Beguá et les *Tšaná-Timbú.

LVIII. Famille Tšibtša (36, I, 29-35, 41-68, 96-142, 143-461 ; 56 *bis*; 61 ; 62).

La famille Tšibtša est une des plus importantes de l'Amérique du Sud. Au Nord, elle empiète largement sur l'Amérique centrale, atteignant la frontière du Costa Rica et du Nicaragua. A l'Ouest, elle a pour limite le Pacifique, sauf au niveau des pays tšoko et esmeralda. Au Sud, dans la région tropicale comprise entre la cordillère et le littoral, elle descend jusqu'à la latitude de Guayaquil. Ses représentants occupent presque tout le haut plateau colombien et une partie du haut plateau équatorien ; enfin, des tribus de même origine se rencontrent sur le versant oriental de la cordillère sur les hauts affluents de l'Orénoque et de l'Amazone.

On divise les dialectes tšibtša en quatre groupes d'après leurs affinités linguistiques :

a) le groupe Talamank-Barbakóa ;

b) le groupe Dorask-Guaymi ;

c) le groupe Tšibtša-Aruak ;
d) le groupe Paez.

a) Le groupe Talamank-Barbakóa comprend :
1) le sous-groupe Guatuso ;
2) le sous-groupe Talamank proprement dit ;
3) le sous-groupe Kuna ;
4) le sous-groupe Barbakóa.

1) Le sous-groupe Guatuso ne comprend que les Guatuso, descendants des anciens *Korobisí, installés sur le río Frio, affluent méridional du San Juan, avec quelques rares représentants sur les ríos Cucaracha, Guacalito, Sapote et sur l'estero Boca Negra.

2) Le sous-groupe Talamank comprend :

α) les *Güetare, autrefois dans les bassins des ríos Grande et Reventazón, et leurs proches parents les *Kepo, entre les ríos Pirris et Grande de Terraba ;

β) les Kabekar, à l'Ouest du río Coén, sur le haut Tarire, et les tribus apparentées des Estrella et des Tširipó sur les ríos de même nom, à l'Ouest de Port Limon, les Tukurriké et Orosi sur le haut Reventazón, les *Suerre sur la côte au Nord des Tširipó ;

γ) les Bribri (Blancos, Biseita, Valientes), au Sud du Coén sur le Lare, affluent du Tariri ;

δ) les Térraba (Tešbi, Tišbi, Depso, Norteños), entre les ríos Tilorio et Tarire, avec la sous-tribu des Tíribi, sur le haut Tilorio ;

ε) les Boruka (Brunka). descendants des anciens *Koto, *Turukaka et *Burukak, dans le bassin du río Grande de Terraba.

3) Le sous-groupe Kuna ne comprend que les Kuna (Kueva, Mandinga, Darien, Tšukunake, Kunakuna, Bayano, Tule, Yule, San Blas). A l'époque de la conquête, ces Indiens s'étendaient à l'Ouest jusqu'à une ligne réunissant l'embouchure du río Cocte sur la côte atlantique au point où le méridien 80° coupe la côte du Pacifique. Au Sud, la limite partait du Pacifique, au niveau de la punta Piñas environ, directement vers l'Est, puis après avoir franchi l'Atrato, gagnait Antioquia sur le Cauca et de là remontait par les sources du río Senu vers la côte atlantique qu'elle atteignait sur la rive orientale du golfe d'Uraba.

4) Au sous-groupe Barbakóa appartiennent :

α) les *Barbakóa, à l'Ouest de la cordillère dans les bassins des ríos Patía, Mira, Cayapas et Esmeraldas et sur le cours supérieur des ríos Daule, Vinces et Bodegas jusqu'au parallèle sud 2° 30', dont les rares survivants sont les Telémbi sur le Patía, les Kuaiker sur le Mira, les Kayápa sur le Cayápas, les Colorado (Saktša, Yumbo) sur les ríos Esmeraldas, Daule et Vinces, et auxquels je rattache les Yumbo, qui vivent à l'Est de Pimampiro ;

β) les *Kara du haut plateau interandin depuis le Chota au Nord jusqu'au parallèle sud 0° 31' ;

γ) les *Kixo (Indiens de la Canela), sur le haut Napo jusqu'à son confluent avec le Coca et sur les rives de cette dernière rivière, et peut-être, les *Latakunga [1] qui vivaient au Sud des Kara.

b) Le groupe Dorask-Guaymi comprend :

1) les *Murire (Bukueta, Sabanero), dans les grandes plaines au Sud de la cordillère, et les vallées profondes du département de Chiriquí jusqu'au río Chame à l'Est ; les *Muoi, dans la vallée de Miranda, sur le río de ce nom ; les *Move (Valientes, Norteños), dans la vallée de Miranda et le long de la côte entre la lagune de Chiriquí et le río Belén, auxquels se rattachent les *Muite du río Coclé del Norte ; les Penonomeños, dans le village de Penonomé ;

2) les *Tšangina (Tšangena) et les *Dorask, dans le bassin du Changuinola, à l'Ouest de la bahia del Almirante, et leurs proches parents, les *Tšumulu de Potrero de Varges, près de Caldera, et les *Gualaka, qui sont des Dorask-Tšangina transplantés dans le département de Chiriquí ;

3) les Tšimila, descendants des anciens *Tairona, dans les forêts entre les pentes occidentales de la Sierra Nevada et le Magdalena.

c) Le groupe Tšibtša-Aruak comprend :

1) les T*šibtša proprement dits, *Muyska ou *Moska (avec la tribu des *Duit, de Duitama), qui vivaient sur le haut plateau colombien entre 4° 15' et 6° 50' de latitude nord, la cordillère orientale à l'Est et le bassin du Magdalena à l'Ouest,

1. Jijón y Caamaño rattache ces deux tribus au groupe Paez.

sans atteindre toutefois les rives du grand fleuve ; les *Kolima (Tapas) et leurs proches parents, les *Muzo, sur la rive droite du Magdalena, entre le Carare au Nord et le río Negro au Sud ;

2) les Rama, qui vivent entre les ríos Bluefields et San Juan, et les Meltšora sur le rio Melchora, derniers représentants des anciens *Voto (Boto), qui habitaient surtout sur les ríos San Carlos et Sarapiqui, affluents du San Juan ;

3) les Aruak, qui se divisent en quatre peuplades, les Kőggaba (Kágaba) sur les pentes septentrionales de la Sierra Nevada dans les villages de San Antonio, San Miguel, Santa Rosa et Pueblo viejo et sur les pentes méridionales dans le village de San José, les Bíntukua (Busintana, Ixka) à San Sebastián, les Guamáka à El Rosario et Marocaso, les Atánkez à Atánquez, tous villages situés également sur le versant sud de la même cordillère ;

4) les Tunebo ou Tame, dont les tribus, Guasiko, Tšita, Morkote, Sínsiga, Tunebo proprement dits et Pedraza s'échelonnent sur les pentes est de la cordillère orientale entre 5° 20′ et 7°, et sur le versant ouest de la cordillère du Cocuy ;

5) les *Betoi et les tribus apparentées des *Xirara, *Situfa, *Ayriko, *Ele, *Lukulia, *Xabúe, *Arauka, *Kilifay, *Anabali, *Lolaka et *Atabaká, à l'Est des Tunebo, sur le haut Casanare, le Cravo, l'Ele, le haut Arauca et le haut Apure ;

6) les Andakí, autrefois installés dans le pays compris entre le Magdalena et le Suaza, notamment dans la vallée de San Agustín, actuellement cantonnés sur le versant oriental de la cordillère, aux sources des deux ríos Fragua.

d) Le groupe Paez comprend :

1) les Paez entre le haut Cauca et le haut Magdalena ; les Panikitá dans le village de ce nom ; les *Kil'a dans la vallée d'Almaguer ;

2) les Totoró dans les villages de Polindará et Totoró ; les Moguex (Guambiano) du village de Silvia ; les Kokonuko des sources du Cauca ; les Guanako dans les hautes montagnes où prend sa source le río Ullucus, affluent du Paez ;

3) les *Pixao, au Nord des Paez, dans le massif montagneux

compris entre le Magdalena à l'Est et le Cauca à l'Ouest; les *Pantše sur les deux rives du Magdalena depuis le Fusagasuga et le Coello au Sud jusqu'au Guali et au río Negro au Nord; les *Kimbaya entre la cordillère centrale à l'Est, la cordillère occidentale à l'Ouest, le Chinchiná au Nord et le río de la Paila au Sud; les *Pantágora sur la rive gauche du Magdalena en aval du Guali.

LIX. Famille Tšikito (**49**, 20-30, 230-242).

Les Tšikito occupent, dans le Sud-Est de la Bolivie, un vaste territoire entre 16° et 18° de latitude et 58° et 62° de longitude, depuis la lagune de Xarayes à l'Est jusqu'au río San Miguel à l'Ouest, avec une tribu isolée à l'Ouest de Santa Cruz, les Tšurápa.

Ils comprennent un nombre considérable de tribus, qui furent réunies dans les missions de San Xavier, Concepción, San Miguel, San Ignacio, Santa Ana, San Rafael, San José, San Juan, San tiago et Santo Corazón.

Les principaux dialectes sont le *Manasika, parlé à proximité des marais Xarayes, aux sources du Paraguay, le Kusikia en usage à Concepción, le Tao dans les environs de San Rafael, le Piñoka près du río San Miguel, le Penoki ou Penokikia à l'Est du río Tunáz, et le Tšurápa sur la rive gauche du río Piray, au Nord-Ouest de Santa Cruz de la Sierra.

LX. Famille *Tširino.

Les *Tširino occupaient les rives du Chirinos, affluent de gauche du Chinchipe, et la région comprise entre cette rivière, le Marañón et la cordillère de Condor.

LXI. Famille Tšoko (**36**, I, 69-95; **65** *bis*).

Les Tšoko habitent le bassin du río Atrato et la côte du Pacifique entre 8° et 4° de latitude nord. Leurs principales tribus portent les noms suivants, empruntés pour la plupart aux villages qu'elles habitent : Chamí, Andagueda ou Angágueda, Murindo, Cañas gordas, Ríoverde, Necoda, Caramanta, Tadocito, Pato, Curusamba, Tucurá, Noanama, Baudocitarae.

LXII. Famille Tšolona.

Les Tšolona vivent sur la rive gauche du Huallaga, sur les affluents Monzon, Uchiza, Tocache et Pachiza, les Tšolona proprement dits entre Tingo-María et El Valle, les Hibito ou Xibito en aval de El Valle.

LXIII. Famille Tšon (9 *bis* ; 37).

Cette famille comprend un groupe Patagon et un groupe Fuégien.

a) Le groupe Patagon comprend :

1) les Tehueltše (Tsoneka), qui habitent la Patagonie du 42e parallèle au détroit de Magellan, et qui comprennent les Peéneken (Paígnk(e)nk(e)n, Pä'änkün'k, Pää'nko-tšönk) et les Ahonikanka (Aóniken, Aônükün'k, Aôniko-tšönk) ;

2) les Téueš (Téhues, Téhuešenk, Téuesson, Tä'uüšn) de la cordillère centrale.

b) Le groupe Fuégien est constitué par les Ona (O'ona, Aona, Aôna), qui occupent toute la Terre de Feu, à l'exception des rives de la baie Useless et du détroit Admiralty, fréquentées par les Alakaluf, et du territoire compris entre le canal de Beagle et la chaîne qui lui est parallèle, où vivent les Yaghan ; ils sont divisés eux-mêmes en Šílk'nam (Šillkanen, Skilkenam^{n}), au Nord de l'île, et en Mánekenkn ou Hauš (Hauss, Haus, Huš, Höš'), à l'Est et au Sud-Est de l'île.

LXIV. Famille Tukáno (Betoya de Brinton) (33 *bis* ; 58 ; 67, 199-201).

La famille Tukáno peut être divisée en trois groupes géographiques : groupe oriental, groupe occidental, groupe septentrional.

a) Le groupe oriental occupe le bassin des ríos Uaupés et Curicuriarý, affluents de droite du río Negro, et de l'Apaporís, affluent de gauche du Yapurá. Les dialectes attestés sont le Tukáno (Daxseá), parlé sur le Curicuriarý, sur l'Uaupés et ses deux branches, le Tiquié et le Caiarý Uaupés ; le Uaíana, le Uásöna, le Kobéua, le Koreá, l'Hölöua, l'Uanána, sur le haut Caiarý Uaupés et ses affluents ; le Karapaná et le Desána, entre le Caiarý Uaupés

et le Papurý ; le Uaikana, sur le bas Papurý ; le Tuyúka, le Bará, le Ömöá, le Buhágana, le Sära et le Tsölá, aux sources du Tiquié ; le Desána, sur le moyen Tiquié ; l'Erúlia, le Palänoa, le Tsöloa, sur le Pirá-paraná, haut affluent de l'Apaporís ; le Dätuana, l'Ópaina ou Tanimbúka [1], le Makúna [1], le Yahúna et le Yupúa, d'amont en aval sur l'Apaporís et ses affluents ; le Kueretú, sur le Miriti-paraná, affluent du Yapurá, et sur le Yapurá même.

b) Le groupe occidental occupe le bassin entier du Napo presque depuis son confluent avec l'Amazone jusqu'à l'embouchure de l'Aguarico, le cours entier de ce dernier, le bassin du Putumayo depuis sa source jusqu'à son confluent avec le río Yaguas, le haut Caquetá jusqu'au 74e degré de longitude environ et son affluent de gauche, le Caguán. La toponymie montre que c'est par cette dernière rivière et le haut Uaupés que s'établissait autrefois la continuité avec le groupe oriental, continuité qui semble avoir été rompue par une migration Karib.

Un très grand nombre de tribus, dont il serait superflu de donner ici la liste, rentre dans ce groupe. Je me contenterai de citer les suivantes : les Tama, qui vivent sur les ríos Yarí et Caguán, affluents de gauche du Yapurá ; les Amaguaxe, sur le haut Caquetá et sur le San Miguel, affluent du Putumayo ; les Koreguaxe, aux sources du Caquetá et du Putumayo ; les Seona ou Zeona, entre le haut Putumayo et le Caquetá ; les Abixira (Avi-širi, Avixira), sur la rive droite du Napo, depuis et y compris le bas Curaray ; les Makaguaxe, sur les ríos Caucaya, Mecaya et Sen-seya, affluents du Caquetá ; les Ikaguate (Ikahuate), Encabellados, Pioxe, Anguteri (Ankotere, Ankutere, Angutero), échelonnés sur le río Aguarico, le cours moyen et inférieur du Napo et entre ce fleuve et le Putumayo.

A ce groupe je rattache, très hypothétiquement et uniquement pour une raison de toponymie, les *Pasto, qui occupaient la région de Pasto dans le Sud de la Colombie et la province équatorienne adjacente du Carchi.

c) Le groupe septentrional comprend exclusivement les Tama et Ayriko, qui vivent aux sources du Manacacia, affluent du Meta.

1. Encore inédit.

LXV. FAMILLE TUPI-GUARANÍ (**17** *bis*, 152-153 ; **21** *bis* ; **30** ; **35**, 445-456 ; **35** *bis*, 59-66 ; **41** ; **42** ; **43** ; **46** *bis* ; **46** *ter* ; **48**, 137-297 ; **49**, 150-197 ; **50**, 331-349 ; **62** *bis* ; **68** ; **68** *bis*, 175-182 ; **76** ; **82**, 307-310).

Le centre de dispersion des Tupi-Guaraní paraît avoir été la région comprise entre le Paraná et le Paraguay. Au moment de la découverte, le Guaraní était la langue dominante dans cette région, c'est-à-dire dans la plus grande partie de l'actuelle République du Paraguay, et les territoires avoisinants, qui correspondent aux provinces argentines de Corrientes, Entre-Ríos, Santa Fe et au gouvernement de Misiones. Parmi les nombreuses tribus guaraní du Paraguay, dont les noms ont été mentionnés, je citerai les *Kariyó (Karió), sur la rive gauche du Paraguay et jusqu'à 100-150 kilomètres à l'intérieur entre le Tibicuary au Sud et le Jejuy au Nord ; les *Paranaé, entre les ríos Tibicuary et Paraná ; les Apapokúva, sur la rive droite du bas Iguatemí à l'extrême Sud de l'État du Matto-Grosso ; les Tañyguá, au Sud des Apapokúva ; les Guayakí ; les Kainguá. La plupart de ces tribus se sont fondues plus ou moins dans la masse de la population du Paraguay moderne, mais, actuellement encore, le Guaraní est la langue dominante dans toute cette république. Pourtant, certains groupements indiens ont encore conservé une indépendance presque complète, tels sont les Guayakí et les Kainguá. Les premiers, restés à l'état primitif, habitent entre 26° et 27° de latitude, entre le Paraná et les sources du Tibicuary, affluent de gauche du Paraguay, sur les pentes sud et sud-est de la sierra de Villa Rica. Les Kainguá (Kaiowa, Kayowa, Kaiguá) sont divisés en deux groupes géographiques : le groupe septentrional occupe le Nord du Paraguay et le Sud du Matto-Grosso, atteignant le río Pardo au Nord et l'Acaray au Sud ; le groupe méridional est installé dans le Sud-Est du Paraguay, aux environs de Jesús et de Trinidad, et dans le territoire argentin adjacent, dans les forêts de San Ignacio et de Corpus (27° de latitude). Les principales tribus Kainguá sont les Apuiteré (Apiteré) entre San Joaquin et le rio Amambahy, et les Tširipá au Nord de

l'Acaray jusqu'au Salto de Guayra. Du groupe septentrional, s'est détachée, au milieu du siècle dernier, une colonie qui s'est installée sur le bas Tibagý, affluent du Paranapanéma, aux environs de São Pedro de Alcantara dans l'État brésilien du Paraná. Les Kainguá des environs de Catanduvas sur l'Iguassú, appelés Batikóla ou Baaberá, et ceux des environs du Salto grande du Paranapanéma et des environs de San Ignacio sur le cours inférieur de la même rivière (Yvytyigua) ont sans doute la même origine.

De même, d'autres tribus guaraní du Paraguay (Apapokúva, Tañyguá et Oguauíva), obéissant à un mouvement migratoire d'ordre religieux vers l'Est, ont remonté, au début du XIX[e] siècle, les affluents de droite du Paraná, ríos Ivahý, Paranapanéma et Tieté, envahi le Brésil méridional et sont venues s'installer entre le haut Aquapehy et les affluents de gauche du Tieté, sur les ríos das Cinzas et Itararé, affluents méridionaux du Paranapanema, et sur la côte entre Santos et Iguapé. Les Aré, qui vivent sur l'Ivahý en aval de Villa Rica, et les Notobotokudo (Pîhtadyovái) des sources de l'Iguassú et de l'Uruguay, qui sont également de parler guaraní, se rattachent peut-être à cette migration, mais ils peuvent appartenir aussi bien aux migrations antérieures, dont je vais parler maintenant.

En effet, avant l'époque historique, des exodes guaraní ont certainement eu lieu; ces migrations, après avoir atteint l'Atlantique, ont suivi le littoral brésilien du Sud au Nord jusqu'à l'embouchure de l'Amazone, et remonté le cours du grand fleuve, presque jusqu'à sa source, et de quelques-uns de ses affluents, particulièrement de ses affluents méridionaux.

Les peuplades tupi-guaraní du littoral ont à peu près entièrement disparu ; il n'en reste actuellement, en dehors des quelques représentants signalés plus haut dans les États de São Paulo et Santa Catharina, que quelques débris parmi la population côtière de Rio grande do Sul, d'Espirito Santo, de Bahia et de Pernambuco. Mais, par les témoignages des anciens auteurs, nous savons qu'au moment de la découverte elles occupaient tout le littoral depuis 30° de latitude jusqu'au bas Amazone. Leurs principales tribus étaient les *Tapé, qui s'étendaient depuis les environs de

Montevideo jusqu'au río Uruguay au Nord, et surtout les *Tupinamba, qui vivaient depuis l'île Santa Catharina au Sud jusqu'à l'embouchure de l'Amazone et sur le cours inférieur de ce fleuve jusqu'au río Negro (l'île de Tupinambarâna garde le souvenir de leur nom), et qui étaient divisés en de nombreuses fractions, les *Kariyó entre l'île Santa Catharina et le río Cananea, les *Tamoyó entre Angra dos Reis et le cap S. Thomé, les *Tupinikiŋ (Tupinaki) entre le río de S. Matheus et Camamú (14°), les *Tupinamba proprement dits entre Camamú et l'embouchure du río San Francisco, les *Potyuára (Pitouara, Potigare, Pitigare, Pitogoare) dans les provinces de Parahyba do Norte et de Ceará jusqu'à la région de Cumá dans l'État de Maranhão, les *Kaité (Kaëté, Kahetê) dans les provinces de Parahyba do Norte, Río Grande do Norte et Ceará, les *Taramembé (Terememembí, Tremembé) entre les ríos Turyassú et Coité (province du Pará), les *Ñengahiba dans l'île de Marajó, etc...

A proximité de la rive méridionale du bas Amazone, il existe encore un certain nombre de tribus tupi-guaraní, qui représentent les restes de cette ancienne migration. Ce sont les Manažé sur le río Ararandéua, affluent du Capím, les Miraño, entre l'Acará et le Capím, aux sources du Bujarú, affluent de gauche de ce dernier; les Tembē sur le río Acará pequeno et le Capím, les Turiwára sur le río Acará grande, les Ararandeũara sur le Mojú, les Anambē sur la rive gauche du bas Tocantins, au-dessous du dernier rapide du Rebojo do Guariba, près de Praia grande (sans doute identiques aux *Amanajo (Manajo) des anciens auteurs) et, apparentés à ceux-ci, les Pakaža (Pakaja), les Žakunda (Jakunda, Amiranha) et les Anta (Tapirauba, Tapirauhũ) à cinq jours de route à l'Ouest de la cataracte Itaboca par 4° de latitude, sur le cours supérieur des rivières qui débouchent dans l'Amazone vers Portel.

Plus à l'Ouest encore, l'Amazone était et est encore en partie occupé par d'autres tribus tupi-guaraní, venues, selon toute vraiemblance, de son cours inférieur. Les *Yurimagua (Zurimagua) vivaient autrefois sur l'Amazone depuis le Purús au moins à l'Est jusqu'au Jutahý à l'Ouest. Fuyant les Portugais, ils émigrèrent

peu à peu vers l'Ouest, en amont du confluent du Putumayo d'abord, puis ultérieurement sur le Huallaga, où un village, situé au confluent de ce fleuve et du Paranapura, porte encore leur nom. Comme eux, les Omagua (Kampeva) viennent de l'Est. Primitivement installés dans les grandes îles du Marañón entre les embouchures du Yuruá et du Napo, ils vinrent se fixer entre le Napo et l'Ucayali, où un village conserve leur nom. Les Kokama (Ukayale) et les Kokamil'a (Gual'aga) firent, sans aucun doute, partie de la même migration. Les premiers vivent sur le bas Ucayali et aux environs de Nauta sur la rive gauche de l'Amazone, les seconds sur le bas Huallaga.

A ces tribus occidentales, mais venus certainement à une époque antérieure, se rattachent les Mirânya (Boro), qui habitent entre le Yapurá et l'Igára-paraná, principalement sur les rives du río Cauinarý, et parlent un dialecte tupi très différencié, alors que l'Omagua et le Kokama sont du Tupi presque pur.

Les Mirânya ne sont pas les seuls représentants de la famille Tupi-Guaraní installés au Nord de l'Amazone. Beaucoup plus à l'Est en effet, on rencontre les Paikipiranga aux sources du Maracá, affluent de gauche du bas Amazone ; les Oyampí (Oyambí, Aiapí, Uajapí, Oaiapí), avec les sous-tribus des Tamakom et des Kussari, qui habitent aux sources de l'Araguary, du Yary et de ses affluents de gauche, dans les monts Tumuc-Humac, et sur l'Oyapok depuis sa source jusqu'au parallèle 3° 30′ environ ; les Emerillon, qui vivent dans la région où prennent naissance l'Approuague, le Camopi, affluent de l'Oyapok, et l'Inini, affluent du Maróni.

Les riverains de l'Amazone ont aussi envahi d'aval en amont un certain nombre des tributaires méridionaux de ce fleuve.

Dans le bassin du Tocantins, on trouve les Kubẽñepre̱ dans les forêts vierges du haut Itacayuna, affluent de gauche en aval du confluent de l'Araguaya, les Tapirapé entre les ríos Tapirapé et Nadjá (entre 10° et 11° de latitude), auxquels se rattachent les Guažažára, des sources de l'Itapucurú et du Mearím ; par contre les Canoeiros (Avá), qui vivent dans le Sud de l'île Bananal et entre le Tocantins et l'Araguaya entre 12° et 14° 30′

de latitude, paraissent être des émigrés venus du Sud au début du XVIIIe siècle (voir p. 692).

Dans le bassin du Xingú, les tribus Tupi-Guaraní signalées jusqu'ici sont les Tekunapéua ou Péua, par 3° 30′ de latitude; les Yuruná, entre 4° 30′ et 8° 30′, et leurs proches parents, les Šipaya (Ašipaye), sur le bas Iriri et son affluent de gauche, le Curuá; les Manitsauá sur la rive gauche vers 11°; les Kamayurá et les Auető sur la rive gauche du Kulisehú, respectivement par 12° 7′ et 12° 23′.

Dans le bassin du río Madeira proprement dit, vivent les Parentintin, Kawahíb ou Kawahíwa, installés entre le Gy-paraná et le Marmellos, avec deux petits groupes isolés sur le Riozinho (Kawahíb-Tupí et Kawahíb-Wiraféd) et le Machadinho, affluents de droite et de gauche du haut Machado. Cette tribu parle un dialecte tupi-guaraní très pur. Dans le même bassin, les Rama-Rama ou Ytangá du río Machadinho, affluent du Machado, et leurs proches parents, les Ntōgapíd du haut Madeirinha, affluent de gauche du río Roosevelt, parlent, au contraire, d'après les courts vocabulaires recueillis par Horta Barbosa et par C. Nimuendajú (ce dernier encore inédit), un dialecte guaraní très altéré, influencé, semble-t-il, par les langues de la famille Katukina.

Dans l'énumération des tribus tupi-guaraní des affluents méridionaux de l'Amazone, j'ai omis de citer jusqu'ici celles du bassin du Tapajoz. Il semble en effet que ces tribus appartiennent, au moins en partie, à un courant migratoire différent de celui que je viens de suivre. Ce second courant, parti directement du Paraguay vers le Nord, aurait envahi le Tapajoz d'amont en aval. Le fait semble prouvé au moins en ce qui concerne les Apiaká, installés sur l'Arinos et le Juruena et au-dessous du confluent de ces deux rivières, et la tribu apparentée des Tapañuma, qui vit sur les deux rives de l'Arinos depuis le 12e parallèle jusqu'à son confluent avec le Juruena; mais de nouvelles études sont nécessaires pour y rattacher les Munduruku (Pari), qui habitent le bas et le moyen Tapajoz depuis le confluent du río S. Manuel, et les Mauhé, installés entre le bas Tapajoz et le bas Madeira, sur le río qui porte leur nom, et dans l'île de Tupinambarâna. La même remarque s'applique à une tribu du Xingú, nettement apparentée

aux Munduruku, les Kuruaya (Kuruahé, Kuruaye, Kurinaye, Kurueye), qui occupent le territoire compris entre l'Iriri et le Curuá, où ils sont parvenus sans doute en suivant le Jamauchim, affluent de droite du Tapajoz. Suivant une tradition très vraisemblable, les Canoeiros du Tocantins seraient aussi venus du Sud ; ils seraient les descendants d'Indiens Kariyó, amenés de São Paulo par les découvreurs de l'État de Goyaz.

Un troisième courant de migration tupi-guaraní s'est dirigé vers l'Ouest.

Les Tširiguano (Aba, Kamba, Tembeta), qui vivent sur les contreforts des Andes boliviennes (serranía de Aguaragüe) et dans les premières plaines de la partie occidentale du Grand Chaco entre 22° et 19° de latitude, depuis les environs de Santa Cruz de la Sierra au Nord jusqu'au haut Bermejo au Sud, sont venus du Paraguay, à travers le Chaco central et méridional, à une date qu'on peut fixer au début du XVIe siècle. Ils trouvèrent dans leur nouvel habitat des tribus d'autre origine, qui adoptèrent leur langue, tels les Tšané ou Tapuí, dont on trouve des restes, d'une part sur le río Itiyuro, près de Campo y Duran, aux sources du Pilcomayo, d'autre part à Caipipendi et sur le río Parapití (où on les appelle Izoceño), chez qui survit, comme langue secrète, un parler arawak. Tels sont également les Tapieté (Tirumbae) ; ces Indiens, qui vivent entre le haut Pilcomayo et le haut Parapití, à l'Est des Tširiguano et au Nord des Tšorotí, entre 21° 30′ et 20°, ont adopté la langue et certaines coutumes des Tširiguano, mais ont une civilisation qui, dans son ensemble, rappelle plutôt celle des Matako, des Tšorotí et des Toba, surtout si on l'observe chez les représentants de la tribu restés à l'état sauvage, les Yanaigua (Yanáygua) du río Parapití.

Les Guarayo, les Siriono et les Pauserna de Bolivie sont venus très probablement aussi du Paraguay, à travers le Chaco boréal. Les Guarayo (Guarayú, Itatin) vivent aux sources du río Blanco et sur le río San Miguel ou Itonama entre 15° et 16° de latitude, les Pauserna sur la rive droite du Guaporé, en amont de son affluent, le Curumbiara, et sur la rive gauche, entre le Paragua et río Verde. Quant aux Siriono, on les rencontre entre le Guaporé

et le bas río Blanco et sur la rive droite de celui-ci vers 14° de latitude et à sa source, entre le haut Machupo et l'Itonama, aux sources de ce dernier fleuve, entre le haut Ivari et le Río Grande, sur le Yapacuní entre le Quimoré et le Piray, et entre ce dernier et le haut Itonama.

La répartition générale des Tupi-Guaraní et l'étude de leur histoire montrent qu'ils ont employé surtout dans leurs migrations la voie maritime et la voie fluviale. C'est pourquoi Hervás les surnommait fort justement les « Phéniciens d'Amérique ».

Une esquisse de grammaire comparée de la famille Tupi-Guaraní a été faite par L. Adam (**1** *bis*). Cet auteur distingue dans le Tupi ancien ou Abañeênga deux dialectes, le dialecte du Sud ou Guaraní proprement dit, le dialecte du Nord ou Tupi proprement dit. Le dialecte du Sud a donné naissance au Guaraní moderne ou Abañeême, tel qu'il est parlé encore au Paraguay et dans les régions avoisinantes, le dialecte du Nord au Tupi moderne ou Ñeêngatu, appelé par les Portugais « lingua geral ». Cette dernière dénomination résulte de ce que le Tupi a eu la même fortune que le Kitšua. Adopté par les missionnaires comme langue d'évangélisation, par les colons comme langue de relation avec les indigènes, le Tupi s'est peu à peu répandu dans tout le bassin de l'Amazone et même dans tout le Brésil, en sorte que presque toutes les tribus indiennes en contact plus ou moins direct avec les Blancs l'ont adopté et sont devenues bilingues, lorsqu'elles n'ont pas oublié complètement leur langue primitive. C'est ainsi que les Katukinarú, établis entre les rivières Embyra et Embyrasu, affluents du Tarauacá, parlent, d'après un vocabulaire recueilli par Bach, un dialecte tupi-guaraní, alors que, selon toute vraisemblance, ils appartiennent à la famille Katukina (**64**) ; de même, le vocabulaire formé par l'abbé E. Ignace chez les Borun (**24**) pourrait nous faire croire que ces Indiens sont des Tupi, alors qu'en réalité ils se rattachent aux Botokudo, par conséquent à la famille Zẹ.

LXVI. Famille Tuyuneiri (47).

Les Tuyuneiri habitent entre le Madre de Dios et ses deux affluents de droite, l'Inambari et le Colorado.

LXVII. Famille Vilela-Tšulupí (Lule de Chamberlain) (**6**, 43-55).

La famille Vilela-Tšulupí occupe dans le Chaco argentin un vaste territoire entre les ríos Bermejo et Salado, du 24ᵉ au 27ᵉ degré de latitude, du 61ᵉ au 65ᵉ degré de longitude.

On y classe :

a) les *Lule (souvent désignés sous le nom de Lule de Machoni), divisés en *Lule proprement dits, *Isistiné, *Tokistiné, *Oristiné et *Tonokoté (auxquels se rattachent les *Matará), dans les plaines des provinces actuelles de Salta, Tucuman et Santiago ;

b) les Vilela, sur les deux rives du Bermejo, divisés en Vilela proprement dits, Tšulupí (Tšunupí, Tšunipí), Pazaine, Sinipé (Sivinipé), Vakaa, Atalalá, Okole, Umuampa, Ipa, Yekoanita, Yook, Teket, Guamaika.

LXVIII. Famille Witóto (**17** *bis*, 148-151 ; **20**, 150-152 ; **32** ; **56** ; **66**, 89-91 ; **67**. 202-205 ; **82**. 299-307).

La famille Witóto comprend :

a) les Witóto proprement dits, qui vivent entre le haut Yapurá et l'Içá, et surtout sur les ríos Carapaná ou Cara-paraná et Igará-paraná, affluents de gauche de l'Içá (tribu des Okaina), dont le dialecte le mieux connu est celui des Witóto-Káimę ;

b) les Miraña-Karapaná-Tapuyo, qui habitent près de la chute Ararakuára du Yapurá, et sont sans doute identiques aux Andoke et aux Nonuya, signalés par les voyageurs modernes ;

c) les Orejones, petite tribu installée sur le río Ambiyacu, affluent de gauche du haut Amazone ;

et avec doute :

d) les *Koëruna, qui vivaient sur la rive gauche du Yapurá vers 2° de latitude sud.

LXIX. Famille Xíbaro (Šiwora, Šuāra) (**17** *bis*, 126-135 ; **28**; **57**).

La famille Xíbaro comprend : les Xíbaro proprement dits et les *Palta.

a) Les Xíbaro proprement dits occupent tout le territoire com-

pris entre la cordillère orientale des Andes à l'Ouest, le río Pastaza au Nord et à l'Est, et le Marañón au Sud, sauf dans la partie comprise entre ses affluents méridionaux, les ríos Nieva et Potro, où une importante tribu, les Aguaruna, occupe sa rive droite.

Ces Indiens se divisent en un très grand nombre de tribus qu'il est inutile d'énumérer ici ; les dialectes suivants sont plus ou moins connus : Makas, Gualakiza, Aguaruna, Zamora, Atšuale, Pintuk, Miazal, Ayuli et Morona.

b) Les *Palta habitaient, dans la vallée interandine équatorienne, une région qui correspond sensiblement à la province actuelle de Loja. Certaines de leurs tribus occupaient également les vallées du haut Zamora et du haut Chinchipe.

LXX. Famille Xiraxara (53).

La famille Xiraxara occupait autrefois, au Vénézuéla, d'une part toute la partie montagneuse de l'État Falcón, la partie septentrionale de l'État Lara et l'Est de l'État Zulia, d'autre part la cordillère côtière à l'Est du Yaracuy, dans les États actuels de Yaracuy, Carabobo et Lara, principalement dans la région de Nirgua.

On y classe les Xiraxara proprement dits, qui habitaient le Nord et l'Ouest du premier centre délimité plus haut et le second centre en entier, et dont les derniers descendants vivent, ou vivaient encore il y a quelques années, à Siquisique ; les Ayamán, riverains du Tocuyo sur son cours moyen, dont les rares survivants se rencontrent dans les environs des municipes de San Miguel, Aguada grande et Moroturo (État Lara) ; les Gayón, cantonnés actuellement dans le municipe de Bobare au Nord-Ouest de Barquisimeto.

Oramas, qui a montré la parenté des langues parlées par ces trois tribus, pense qu'il faut les rattacher à la famille Arawak, mais les preuves qu'il a données à l'appui de cette hypothèse ne sont pas concluantes.

LXXI. Famille Yahgan (9 *bis*).

Les Yahgan (Yámana) occupent ou occupaient la côte méri-

dionale de la Terre de Feu, d'une extrémité à l'autre, et tout l'archipel méridional : îles Navarin, Cap Horn, Hoste, Gordon, Londonderry, Basket, etc...

LXXII. Famille Yaruro.

Les Yaruro (Pumé, Yuapín) vivent dans le bassin du Capanaparo, affluent de gauche de l'Orénoque.

LXXIII. Famille *Yunka.

Les *Yunka occupaient les vallées chaudes du Pérou, près de la mer, entre 6° 30′ et 10° de latitude méridionale, leur centre principal étant la région de la ville actuelle de Trujillo. On classe dans cette famille les tribus suivantes, en allant du Nord au Sud : les *Morrope, au Nord de Lambayeque, les *Eten, au Sud de Lambayeque, les *Tšimu, près de Trujillo, les *Motšika ou *Tšintša, à Moche, au Sud de Trujillo, les *Tšanko, au Sud des Motšika.

LXXIV. Famille Yurakáre (50, 564-568).

Les Yurakáre (Yuruxure, Yurukare) vivent entre les parallèles 16° et 17°, aux sources du Sécure, du Chapáre et du Chimoré, affluents de gauche du Mamoré.

LXXV. Famille Yuri (64).

Les Yuri vivaient sur le bas Yapurá, depuis le delta jusqu'aux premières chutes, et entre ce fleuve et l'Içá ou Putumayo, principalement sur le río Puré supérieur, vers le parallèle sud 2° 30′.

LXXVI. Famille Záparo (57).

Les Záparo, dont on sait les noms de 39 tribus, vivent entre le Napo, le Bobonaza et le Pastaza. On connaît cinq dialectes de leur langue : le Záparo proprement dit, le Konambo, le Gae, l'Andoa et l'Ikito. Ils présentent des ressemblances lexicographiques assez nettes avec le Mirânya, qui est un dialecte tupiguaraní très différencié, mais sans qu'on puisse en conclure à une parenté originelle.

LXXVII. FAMILLE Žẹ (Gẹ) ou KRAN[1] (Tapuyo de Brinton et Chamberlain) (3 ; 7 ; 25 ; 26; 29 ; 34 ; 35 *ter* ; 45 ; 46 ; 51 ; 70 ; 72 ; 79).

La famille Žẹ occupe toute la moitié orientale du plateau brésilien, depuis sa chute septentrionale marquée par les dernières cataractes du Xingú et du Tocantins (3° 30′ de latitude sud environ) jusqu'au 30ᵉ parallèle au Sud, et au haut Xingú à l'Ouest ; en aucun point, elle n'atteint l'Amazone.

On divise les nombreuses tribus de parler Žẹ en quatre groupes : les Žẹ orientaux, les Žẹ septentrionaux, les Žẹ centraux et les Žẹ méridionaux.

a) Les Žẹ orientaux vivent dans les forêts vierges des pentes orientales du plateau brésilien et sur les fleuves côtiers qui y naissent, depuis le río Pardo au Nord jusqu'au río Doce au Sud. Ils comprennent :

1) les Botokudo (Burung, Borung, Borun), descendants des anciens *Aimoré, qui habitent l'État d'Espirito Santo, la partie orientale de l'État de Minas Geraes et la partie méridionale de l'État de Bahia, notamment dans les bassins des ríos Doce, Mucury et Belmonte, jusqu'au río Pardo au Nord, et au río Preto, affluent du Parahyba, au Sud (tribu des *Ararýs) ;

2) les Kamakan (Mongoyó ou Monšokó), entre les ríos Pardo, Ilheos et de Contas, dans le Sud de l'État de Bahia, et les tribus apparentées des *Katathoý, des *Kotošó (Kutašó, Katašó), dans les montagnes entre les ríos de Contas et Pardo, des Menien (Menian ou Menieng), sur le bas río Belmonte ou Jequitinhonha, et des Masakará, dont Martius rencontra les derniers représentants à Joazeiro, sur le río San Francisco ;

3) les *Pañame, sur le Sussuhý pequeno, affluent septentrional du río Doce ; les *Kapošó et les *Kumanašó, dans les montagnes entre Minas Geraës et Porto Seguro ; les *Mašakalí, sur le río Belmonte, près de S. Miguel, associés aux *Patašó entre le río de

1. Cette famille est, de toutes les familles sud-américaines, celle qui a été le plus artificiellement constituée. Elle est comme le *caput mortuum* de la linguistique sud-américaine. Sa révision soigneuse et complète s'impose, sur des bases vraiment scientifiques.

Santa Cruz et le río de S. Matheus; les *Makuní et les *Monošó, autrefois dans les montagnes aux limites de Minas, Porto Seguro et Bahia, puis réunis à Alto dos Boys dans le district de Minas novas ; les *Malalí, sur le Sussuhý pequeno, et les *Koropó, sur le río da Pomba [1];

4) les *Coroados et les *Purí, qui vivaient depuis le río Parahyba jusqu'au río de Espirito Santo et pénétraient dans la partie orientale de Minas Geraes; les *Goyataká, entre le bas Parahyba et le Macahé [2].

b) Les Žę septentrionaux comprennent les peuplades de l'État de Maranhão, connues sous le nom général de Timbira, qu'on divise en Timbira de forêt et en Timbira de savane.

1) Les Timbira de forêt comprennent les Timbira proprement dits, sur le haut Gurupy; les Gamella ou Akobū, sur le moyen Itapicurú; les Sakamekran, à l'Ouest du río Itapicurú entre Caxias et Pastos bons sur le cours supérieur du Codó et du Mearim; les Piokobžę, sur le Grajaú; les Augutžę, les Kranžę (dont les derniers survivants, les Krenžę ou Tājė, vivent sur la rive droite du bas Mearim) et les Paikožę, à la frontière des États de Marañhao et du Pará entre 3° et 5° de latitude.

2) Les Timbira de savane ou Canellas finas comprennent les Temembū, Bukobū ou Mannožó (sans doute identiques aux Aponegikran), sur la rive droite du bas Tocantins entre 3° 30′ et 5° de latitude sud; les Kapiekran, autrefois aux sources du río das Balsas, actuellement dans divers villages près de la serra dos Canelas et dans les missions de Grajahú et Barra do Corda; les Makamekran, appelés Karaôu ou Kraô dans le territoire de Pastos bons, Tamembó, Pepuši et Karahú (Karakú) sur la rive droite du Tocantins (entre Boa vista et l'embouchure de l'Araguaya), autrefois répandus dans tout le territoire compris entre le Parna-

1. Ces tribus, à l'exception des Malalí, composaient la famille Goyataká de Martius. Ils forment un groupe de dialectes affines, dont la parenté avec les autres parlers de la famille Žę aurait besoin d'être confirmée.

2. Ces trois tribus avec les Koropó (que je crois plus juste de classer dans le sous-groupe précédent) constituent la famille Goytacaz (Goyataka, Waitakka) d'Ehrenreich et de Chamberlain. Leur rattachement à la famille Zę, admis par d'autres auteurs, demande confirmation.

hyba, le río das Balsas, le río Manuel Alves grande et le Tocantins, actuellement réduits à une petite tribu aux sources du río Preto, affluent du río Grande; les Kannakažę (Kannakatžę), dans la vallée du río Farinha; les Purekamekran, en aval de ce río sur la rive droite du Tocantins; les Ponkatžę et les Norokoažę, sur la même rive, en aval des précédents.

c) Les Žę centraux comprennent le sous-groupe Kayapó et le sous-groupe Akuä.

1) Au début du XVIIIe siècle, les Kayapó peuplaient le Goyaz méridional, d'où ils émigrèrent en partie vers le Sud, sur le Paranahyba, affluent du haut Paraná, — où ils étaient arrivés au début du XIXe siècle —, qu'ils occupent depuis le río Anicuns jusqu'au río Pardo. Les Kayapó restés dans l'État de Goyaz furent en partie réduits à Mossamedes, cependant que la plupart s'unissaient avec les tribus apparentées de la rive occidentale du moyen Araguaya, où les Mäkubengokrę́-Kayapó, les Gradahō (Gradaú), les Kurupite (probablement identiques aux Purukarú de Coudreau), les Gorotiré, les Karaho et les Ušikring (Chicrî) vivent encore en pleine indépendance, entre les 7e et 10e parallèles. Les plus occidentaux de ces Kayapó du Nord sont les Suyá de la rive droite du haut Xingú, en amont de la cataracte Martius, vers 11° 15′, les plus septentrionaux, les Apinažę, établis sur le Tocantins et l'Araguaya en amont et près de leur confluent, et les Gaviões ou Krikatažę ou Karakati, qui vivaient sur le premier de ces fleuves vers le 5e degré de latitude.

2) Le sous-groupe Akuä comprend les Šavánte Opaié (Araē), qui vivent entre le Tocantins et l'Araguaya depuis Boa vista au Nord jusqu'au 12e parallèle au Sud, et sur la rive gauche de l'Araguaya dans le bassin du río dos Mortes; les Šerénte, autrefois sur la rive droite du Tocantins depuis le río Manuel Alves grande au Nord jusqu'au río do Peixe au Sud, actuellement cantonnés entre Boa vista et le río de Somno ; les *Šikriabá, autrefois sur le río Preto, affluent du río Grande, avec une colonie aux sources du Paranahyba, origine du Paraná; les *Žeiko (Geiko, Jeiko, Jaiko) autrefois sur les ríos Canindé et Gurguêa et sur la ligne de partage des eaux entre ces deux fleuves et le río San Francisco; les

*Akroá, autrefois entre le río San Francisco et le Tocantins depuis les sources du Parnahyba (dans l'État de Piauhy) jusqu'aux sources du Paranahyba, origine du Paraná, dont une tribu, les *Goguē ou *Gueguē, vivait entre le Tocantins, le río de Somno et le Gurguêa.

d) Les Žę méridionaux ou Guayanã occupaient au XVI[e] siècle le territoire qui s'étend depuis Rio grande do Sul et le Nord de l'Argentine, à travers les États de São Paulo et de Rio de Janeiro, jusqu'à l'intérieur de l'État de Bahia. Leurs descendants sont divisés actuellement en deux sous-groupes : un sous-groupe oriental et un sous-groupe occidental.

1) Le sous-groupe oriental est constitué par les Kaingáng [1] (Kaingán, Kaingýgn, Bugre, Kamé), établis en Argentine dans le Gouvernement de Misiones, sur un petit affluent du río Yaboti à 5 kilomètres de San Pedro del Monte, et au Brésil dans la partie occidentale des États de Rio grande do Sul, Santa Catharina, Paraná et São Paulo, à l'Ouest de la serra do Mar, sur le Paranapanéma, le Tibagý, l'Ivahý, le Piquirý et l'Iguassú, affluents du Paraná, et même sur les affluents supérieurs de l'Uruguay. L'appellation de Bugre s'applique plus particulièrement à la tribu des Šokren (Šokleng, Šokré), qui vit dans les forêts vierges de l'État de Santa Catharina entre l'Iguassú et le Piquirý et aux sources de l'Uruguay ; l'appellation de Guayañã (anciennement Guanhanan) à la tribu installée dans l'Ouest de l'État de São Paulo, dans les municipes d'Itapéva et de Faxina ; l'appellation de Kamé à la tribu qui vit sur le littoral de São Paulo, entre Santos et Iguape, et aux Indiens des environs de Guarapuava (Paraná) et de Palmas (Santa Catharina). Les Kaingáng de la rive droite du Paranapanéma s'appellent Nhakfáteitei.

2) Le sous-groupe occidental est formé par les Ingain, installés sur les deux rives du haut Paraná, entre le torrent Iuitorocay et la chute de Guayra, et par la tribu, étroitement apparentée, des Guayaná (Waiganna), signalée dans la même région et un peu plus au Sud, aux environs d'Azara, sur la rive droite du Paraná.

1. Appelés Coroados par les Brésiliens, Tupí par les Argentins.

INDEX BIBLIOGRAPHIQUE.

1. Adam (Lucien). *Matériaux pour servir à l'établissement d'une grammaire comparée des dialectes de la famille Caribe*. Bibliothèque linguistique américaine, t. XVII. Paris, 1893.

1 *bis*. Adam (Lucien). *Matériaux pour servir à l'établissement d'une grammaire comparée des dialectes de la famille Tupi*. Bibliothèque linguistique américaine, t. XVIII. Paris, 1896.

2. Alvarado (Lisandro). *Observaciones sobre el Caribe hablado en los llanos de Barcelona*. Caracas, 1919.

2 *bis*. Augusta (Félix José de). *Diccionario Araucano-Español y Español-Araucano*. Santiago de Chile, 2 vol., 1916.

2 *ter*. Bach (J.). *Datos sobre los Indios Terenas de Miranda*. Anales de la Sociedad científica argentina. Buenos Aires, t. LXXXII, 1916, p. 87-94.

3. Barcatta de Val Floriana (Mansueto). *Diccionarios Kainjgang-Portuguez e Portuguez-Kainjgang*. Revista do Museu paulista. S. Paulo, t. XII, 1920.

4. Bibolotti (Benigno). *Moseteno vocabulary and treatises* (with an introduction by Rudolph Schuller). Evanston et Chicago, 1917.

5. Bolinder (Gustaf). *Einiges über die Motilon-Indianer der Sierra de Perijá*. Zeitschrift für Ethnologie. Berlin, t. XLIX, 1917, p. 21-51.

6. Boman (Éric). *Antiquités de la région andine de la République argentine et du désert d'Atacama*. Mission scientifique G. de Créqui-Montfort et E. Sénéchal de la Grange. Paris, 2 vol., 1908.

7. Bruno (Rudolph). *Wörterbuch der Botokudensprache*. Hambourg, 1909.

8. Buchwald (Otto von). *Migraciones sud-americanas*. Boletín de la Sociedad ecuatoriana de estudios históricos americanos. Quito, t. I, 1918, p. 227-236.

8 *bis*. Buchwald (Otto von). *El Sebondoy ; vocabulario y notas*. Boletín de la Sociedad ecuatoriana de estudios históricos americanos. Quito, t. III, 1919, p. 205-212.

8 *ter*. Cañas Pinochet (Alejandro). *Estudios de la lengua Veliche*. Trabajos del cuarto Congreso científico (1° pan-americano) celebrado en Santiago de Chile del 25 de diciembre de 1908 al 5 de enero de 1909, t. XI. Santiago de Chile, 1911, p. 143-330.

9. Chamberlain (Alexander F.). *Linguistic stocks of South American Indians, with distribution-map*. American anthropologist. Lancaster, new series, t. XV, 1913, p. 236-247.

9 *bis*. Cooper (John M.). *Analytical and critical bibliography of the tribes of Tierra del fuego and adjacent territory*. Bureau of american ethnology, Bulletin 63. Washington, 1917.

9 *ter*. Coryn (Alfredo). *Los Indios Lenguas, sus costumbres y su idioma con compendio de gramática y vocabulario*. Anales de la Sociedad científica argentina. Buenos Aires, t. XCIII, 1922, p. 221-282.

10. Créqui-Montfort (G. de) et Rivet (P.). *Linguistique bolivienne. La*

langue Kaničana. Mémoires de la Société de linguistique de Paris, t. XVIII, 1914, p. 354-377.

11. CRÉQUI-MONTFORT (G. de) et RIVET (P.). *Les affinités des dialectes Otukè*. Journal de la Société des Américanistes de Paris, nouv. série, t. X, 1913, p. 369-377.

12. CRÉQUI-MONTFORT (G. de) et RIVET (P.). *Linguistique bolivienne. La langue Mobima*. Journal de la Société des Américanistes de Paris, nouv. série, t. XI, 1914-1919, p. 183-211.

13. CRÉQUI-MONTFORT (G. de) et RIVET (P.). *Linguistique bolivienne. La langue Kayuvava*. International Journal of american linguistics. New York, t. I, 1917-1920, p. 245-265.

14. CRÉQUI-MONTFORT (G. de) et RIVET (P.). *Linguistique bolivienne. La langue Uru ou Pukina*. Internationales Archiv für Ethnographie. Leyde, t. XXV, fasc. 3-4, 1921, p. 87-113.

15. CRÉQUI-MONTFORT (G. de) et RIVET (P.). *La famille linguistique Takana*. Journal de la Société des Américanistes de Paris, nouv. série, t. XIII, 1921, p. 91-102, 281-301 ; t. XIV, 1922, p. 141-182 ; t. XV, 1923, p. 121-167.

16. EHRENREICH (Paul). *Die Ethnographie Südamerikas im Beginn des XX. Jahrhunderts unter besonderer Berücksichtigung der Naturvölker*. Archiv für Anthropologie. Braunschweig, neue Folge, t. III, 1905, p. 59-75.

17. FARABEE (William Curtis). *The central Arawaks*. University of Pennsylvania, The University Museum, Anthropological Publications. Philadelphie, t. IX, 1918.

17 *bis*. FARABEE (William Curtis). *Indian tribes of eastern Peru* (introduction of Louis John de MILHAU). Papers of the Peabody Museum of american archaeology and ethnology, Harvard University. Cambridge, t. X, 1922.

18. GOEJE (C. H. de). *Études linguistiques caraïbes*. Verhandelingen der koninklijke Akademie van Wetenschappen te Amsterdam. Afdeeling Letterkunde. Amsterdam, nieuwe Reeks, t. X, nº 3, janvier 1910.

19. GRUBB (W. Barbrook). *An unknown people in an unknown land ; an account of the life and customs of the Lengua Indians of the paraguayan Chaco, with adventures and experiences met with during twenty years' pioneering and exploration amongst them*. Londres, 1911.

20. HARDENBURG (W. E.). *The Putumayo, the devil's paradise ; travels in the peruvian Amazon region and an account of the atrocities committed upon the Indians therein*. Londres, 1912.

21. HESTERMANN (Ferd.). *Die Schreibweise der Pano-Vokabularien*. Journal de la Société des Américanistes de Paris, nouv. série, t. XI, 1914-1919, p. 21-33.

21 *bis*. HORTA BARBOSA (Nicolau Bueno). *Exploração e levantamento dos rios Anary e Machadinho*. Commissão de linhas telegraphicas estrategicas de Matto-Grosso ao Amazonas, Publicação n. 48. Rio de Janeiro, [1922].

22. HUNT (Richard J.). *El Vejoz ó Aiyo* (Introducción de Samuel A. LAFONE-QUEVEDO). Revista del Museo de La Plata. Buenos Aires, t. XXII, 1913, p. 7-215.

23. HUNT (Richard J.). *El Choroti ó Yófuaha* (Introducción de Samuel A. LAFONE-QUEVEDO). Revista del Museo de La Plata. Liverpool, t. XXIII, 1re partie, 1915.

24. IGNACE (Étienne). *Les Boruns.* Anthropos. St.-Gabriel-Mödling, t. IV, 1909, p. 942-944.

25. IGNACE (Étienne). *Les Camacans.* Anthropos. St.-Gabriel-Mödling, t. VII, 1912, p. 948-956.

26. IHERING (Hermann von). *A ethnographia do Brasil meridional.* Actas del XVII° Congreso internacional de Americanistas, sesión de Buenos Aires, 17-23 de mayo de 1910. Buenos Aires, 1912, p. 250-264.

27. JAHN (A.). *Paraubanos und Guajiros und die Pfahlbauten am See von Maracaibo.* Zeitschrift für Ethnologie. Berlin, t. XLVI, 1914, p. 267-283, 536.

28. KARSTEN (Rafael). *La lengua de los Indios Jíbaros (Shuāra) del Oriente del Ecuador. Gramática, vocabulario y muestras de la prosa y poesía.* Finska Vetenskaps-Societetens Förhandlingar. Helsingfors, t. LXIV, 1921-1922, Avd. B, n° 2.

28 *bis.* KARSTEN (Rafael). *The Toba Indians of the Bolivian Gran Chaco.* Acta Academiae aboensis Humaniora IV. Åbo, 1923.

29. KISSENBERTH (Wilhelm). *Bei den Canella-Indianern in Zentral-Maranhão (Brasilien).* Baessler-Archiv. Berlin, t. II, 1912, p. 45-54.

30. KISSENBERTH (Wilhelm). *Beitrag zur Kenntnis der Tapirapé-Indianer.* Baessler-Archiv. Berlin, t. VI, 1922, p. 36-81.

31. KOCH (Theodor). *Die Maskoi-Gruppe im Gran Chaco.* Mittheilungen der anthropologischen Gesellschaft in Wien, t. XXXII, 1902, p. 130-148.

32. KOCH-GRÜNBERG (Theodor). *Die Uitóto-Indianer. Weitere Beiträge zu ihrer Sprache* (nach einer Wörterliste von Hermann SCHMIDT). Journal de la Société des Américanistes de Paris, nouv. série, t. VII, 1910, p. 61-83.

33. KOCH-GRÜNBERG (Theodor). *Abschluss meiner Reise durch Nordbrasilien zum Orinoco, mit besonderer Berücksichtigung der von mir besuchten Indianerstämme.* Zeitschrift für Ethnologie. Berlin, t. XLV, 1913, p. 448-474.

33 *bis.* KOK (P. P.). *Ensayo de gramática Dagseje o Tokano.* Anthropos. St.-Gabriel-Mödling, t. XVI-XVII, 1921-1922, p. 838-865.

34. KRAUSE (Fritz). *In den Wildnissen Brasiliens.* Leipzig, 1911.

34 *bis.* KUNIKE (Hugo). *Die Phonetik der Karaiásprache (nach linguistischen Principien).* Journal de la Société des Américanistes de Paris, nouv. série, t. XI, 1914-1919, p. 139-181.

34 *ter.* KUNIKE (Hugo). *Beiträge zur Phonetik der Karajá-Sprache (Brasilien).* Internationales Archiv für Ethnographie. Leyde, t. XXIII, 1916, p. 147-182.

35. LANGE (Algot). *The lower Amazon. A narrative of explorations in the little known regions of the state of Pará, on the lower Amazon, with a record of archaeological excavations on Marajó island at the mouth of the Amazon river, and observations on the general resources of the country.* New York, 1914.

35 *bis.* LEAL (Oscar). *O Amazonas.* Lisbonne, 1894.

35 *ter.* LEAL (Oscar). *Viagem a um paiz de selvagens.* Lisbonne, 1895.

36. LEHMANN (Walther). *Zentral-Amerika*. I Teil. *Die Sprachen Zentral-Amerikas*. Berlin, 2 vol., 1920.

37. LEHMANN-NITSCHE (R.). *El grupo lingüistico Tshon de los territorios magallánicos*. Revista del Museo de La Plata. Buenos Aires, t. XXII, 1913, p. 217-276.

38. LEHMANN-NITSCHE (R.). *El grupo lingüistico « Het » de la pampa argentina*. Revista del Museo de La Plata. Buenos Aires, t. XXVII, 1923, p. 10-85.

39. LEHMANN-NITSCHE (R.). *El grupo lingüistico Alakaluf de los canales magallánicos*. Revista del Museo de La Plata. Buenos Aires, t. XXV, 1919, p. 15-69.

40. MAGALHÃES (Basilio de). *Vocabulario da lingua dos Boróros-Coroados do Estado de Matto-Grosso*. Revista do Instituto historico e geographico brasileiro. Rio de Janeiro, t. LXXXIII, 1918 (1919), p. 5-67.

41. MAYNTZHUSEN (F. C.). *Die Sprache der Guayaki*. Zeitschrift für Eingeborenensprachen. Berlin, t. X, 1919-1920, fasc. 1.

42. NIMUENDAJÚ-UNKEL (Curt). *Die Sagen von der Erschaffung und Vernichtung der Welt als Grundlagen der Religion der Apapocúva-Guarani*. Zeitschrift für Ethnologie. Berlin, t. XLVI, 1914, p. 284-403.

43. NIMUENDAJÚ-UNKEL (Curt). *Vocabularios da lingua geral do Brazil nos dialectos dos Manajé do rio Ararandéua, Tembé do rio Acará pequeno, é Turiwára do rio Acará grande, Est. do Pará*. Zeitschrift für Ethnologie. Berlin, t. XLVI, 1914, p. 615-618.

44. NIMUENDAJÚ-UNKEL (Curt). *Vokabular der Pariri-Sprache*. Zeitschrift für Ethnologie. Berlin, t. XLVI, 1914, p. 619-625.

45. NIMUENDAJÚ-UNKEL (Curt). *Vokabular und Sagen der Crengêz-Indianer (Tājê)*. Zeitschrift für Ethnologie. Berlin, t. XLVI, 1914, p. 626-636.

46. NIMUENDAJÚ-UNKEL (Curt). *Vocabulare der Timbiras von Maranhão und Pará*. Zeitschrift für Ethnologie. Berlin, t. XLVII, 1915, p. 302-305.

46 *bis*. NIMUENDAJÚ (Curt). *Os Indios Parintintin do rio Madeira*. Journal de la Société des Américanistes de Paris, nouv. série, t. XVI, 1924.

46 *ter*. NIMUENDAJÚ (Curt) et BENTES (E. H. do Valle). *Documents sur quelques langues peu connues de l'Amazone*. Journal de la Société des Américanistes de Paris, nouv. série, t. XV, 1923, p. 215-222.

47. NORDENSKIÖLD (Erland). *Beiträge zur Kenntnis einiger Indianerstämme des Rio Madre de Dios-Gebietes*. Ymer. Stockholm, 1905, p. 265-312.

48. NORDENSKIÖLD (Erland). *Indian life i Gran Chaco (Syd-Amerika)*. Stockholm, 1910.

49. NORDENSKIÖLD (Erland). *Indianer och Hvita i nordöstra Bolivia*. Stockholm, 1911.

50. NORDENSKIÖLD (Erland). *Forskningar och äventyr i Sydamerika 1913-1914*. Stockholm, 1915.

51. OLIVEIRA (J. Feliciano de). *The Cherentes of central Brazil. Notes for a sketch*. International Congress of Americanists, Proceedings of the XVIII session, Londres, 1912. Londres, 1913, part II, p. 391-396.

52. ORAMAS (Luis R.). *Contribución al estudio de los dialectos Puinabe y Maquiritare.* Gaceta de los Museos nacionales. Caracas, t. I, 1912-1913, p. 20-27.

52 *bis*. ORAMAS (Luis R.). *Contribución al estudio de la lengua Guajira.* Publicaciones de la « Revista técnica del Ministerio de obras públicas » de los EE. UU. de Venezuela. Caracas, 1913.

53. ORAMAS (Luis R.). *Materiales para el estudio de los dialectos Ayamán, Gayón, Jirajara, Ajagua.* Caracas, 1916.

54. ORAMAS (Luis R.). *Estudios lingüísticos. Patronímicos Quiriquires y vocabulario Paraujano comparado con el Guajiro.* De Re indica. Caracas, t. I, nº 1, sept. 1918, p. 23-28 ; nº 2, oct. 1918, p. 43.

55. OUTES (Félix F.). *Sobre las lenguas indígenas rioplatenses.* Revista de la Universidad de Buenos Aires, t. XXIV, 1913, p. 231-237.

56. PREUSS (Konrad Theodor). *Religion und Mythologie der Uitoto.* Göttingen, t. I, 1921 ; t. II, 1923.

56 *bis*. PREUSS (Konrad Theodor). *Forschungsreise zu den Kágaba-Indianern der Sierra Nevada de Santa Marta in Kolumbien.* Anthropos. St.-Gabriel-Mödling, t. XIV-XV, 1919-1920, p. 314-404, 1040-1079 ; t. XVI-XVII, 1921-1922, p. 459-480, 737-764.

57. RIVET (P.). *Les familles linguistiques du Nord-Ouest de l'Amérique du Sud.* L'Année linguistique. Paris, t. IV, 1908-1910 (1911), p. 117-154.

58. RIVET (P.). *La famille Betoya ou Tukano (Note complémentaire).* Mémoires de la Société de linguistique de Paris, t. XIX, 1916, p. 91-95.

59. RIVET (P.). *Affinités du Sáliba et du Piaróa.* Journal de la Société des Américanistes de Paris, nouv. série, t. XII, 1920, p. 11-20.

60. RIVET (P.). *Nouvelle contribution à l'étude de la langue des Itonama.* Journal de la Société des Américanistes de Paris, nouv. série, t. XIII, 1921, p. 173-195.

61. RIVET (P.). *La langue Andaki.* Journal de la Société des Américanistes de Paris, nouv. série, t. XVI, 1924, p. 99-110.

62. RIVET (P.). *La langue Tunebo.* Journal de la Société des Américanistes de Paris, nouv. série, t. XVI, 1924, p. 19-92.

62 *bis*. RIVET (P.). *Les Indiens Canoeiros.* Journal de la Société des Américanistes de Paris, nouv. série, t. XVI, 1924, p. 169-181.

63. RIVET (P.) et TASTEVIN (C.). *Affinités du Makú et du Puinâve.* Journal de la Société des Américanistes de Paris, nouv. série, t. XII, 1920, p. 69-82.

64. RIVET (P.) et TASTEVIN (C.). *Les tribus indiennes des bassins du Purús, du Juruá et des régions limitrophes.* La Géographie. Paris, t. XXXV, 1921, p. 449-482.

65. RIVET (P.) et TASTEVIN (C.). *Les langues du Purús, du Juruá, et des régions limitrophes. 1º. Le groupe arawak pré-andin.* Anthropos. St.-Gabriel-Mödling, t. XIV-XV, 1919-1920, p. 857-890 ; t. XVI-XVII, 1921-1922, p. 298-325, 819-828 ; t. XVIII-XIX, 1923-1924, p. 104-113.

65 *bis*. ROBLEDO (Emilio). *Vocabulario de los Chamíes.* Repertorio histórico. Medellín, 4e année, 1922, p. 603-607.

66. ROBUCHON (Eugenio). *En el Putumayo y sus afluentes*. Lima, 1907.

67. ROCHA (Joaquin). *Memorandum de viaje (Regiones amazónicas)*. Bogotá, 1905.

68. ROMANO (P. Santiago) et CATTUNAR (P. Herman). *Diccionario Chiriguano-Español y Español-Chiriguano, compilado teniendo á la vista diversos manúscritos de antiguos misioneros del apostólico Colegio de Santa Maria de los Angeles de Tarijá y particularmente el diccionario Chiriguano etimológico del R. P. Doroteo Giannecchini*. Tarija, 1916.

68 *bis*. RONDON (Candido Mariano da Silva). *Relatorio apresentado á Divisão de engenharia (G. 5) do Departamento da guerra e á Directoria geral dos telegraphos*. Commissão de linhas telegraphicas estrategicas de Matto-Grosso ao Amazonas, Publicação n. 26. Rio de Janeiro, 1915.

69. ROQUETTE-PINTO (E.). *Rondonia. Anthropologia, Ethnographia*. Archivos do Museu nacional do Rio de Janeiro, t. XX, 1917.

70. SALA (Antonio Maria). *Ensaio de grammatica Kaiapó*. Revista do Museu paulista. S. Paulo, t. XII, 1920.

71. SALA (R. P. Fray Gabriel). *Diccionario, gramática y catecismo Castellano, Inga, Amueixa y Campa*. Boletín de la Sociedad geográfica de Lima, t. XVII, 1905, p. 149-227, 313-356, 471-490 ; t. XIX, 1906, p. 102-120.

72. SAMPAIO (Theodoro). *Os Kraôs do Rio Preto no Estado da Bahia*. Revista do Instituto historico e geographico brasileiro. Rio de Janeiro, t. LXXV, 1912 (1913), parte I, p. 143-205.

72 *bis*. SCHMIDT (Max). *Die Paressi-Kabiši. Ethnologische Ergebnisse der Expedition zu den Quellen des Jauru und Juruena im Jahre 1910*. Baessler-Archiv. Berlin, t. IV, 1914, p. 167-250.

73. SCHMIDT (Max). *Die Guato und ihr Gebiet. Ethnologische und archäologische Ergebnisse der Expedition zum Caracara-Fluss im Matto-Grosso*. Baessler-Archiv. Berlin, t. IV, 1914, p. 251-283.

74. SCHULLER (Rudolph R.). *Discovery of a fragment of the printed copy of the work on the Millcayac language by Luis de Valdivia with a bibliographical notice*. Papers of the Peabody Museum of american archaeology and ethnology, Harvard University. Cambridge, t. III, 1904-1913, p. 221-258.

75. SCHULLER (Rudolph R.). *The only known words of the Charrúa language of Rio de la Plata*. Proceedings of the nineteenth international Congress of Americanists, Washington, 1915. Washington, 1917, p. 552-554.

76. SNETHLAGE (E.). *Zur Ethnographie der Chipaya und Curuahé*. Zeitschrift für Ethnologie. Berlin, t. XLII, 1910, p. 612-637.

76 *bis*. SOUSA (Antonio Pyrineus de). *Exploração do rio Paranatinga e seu levantamento topographico bem como o dos rios S. Manoel e Telles Pires*. Commissão de linhas telegraphicas estrategicas de Matto-Grosso ao Amazonas, Publicação n. 34. Rio de Janeiro, 1916.

77. TAVERA-ACOSTA (B.). *Nuevos vocabularios de dialectos indigenas de Venezuela*. Journal de la Société des Américanistes de Paris, nouv. série, t. XIII, 1921, p. 217-232 ; t. XIV, 1922, p. 65-82.

78. TELLO (Julio C.). *Arawak (Fragmento de linguistica indigena sudamericana)*. Lima, 1913.

79. TESCHAUER (P. C.). *Die Caingang oder Coroados-Indianer im brasilianischen Staate Rio Grande do Sul*. Anthropos. St.-Gabriel-Mödling, t. IX, 1914, p. 16-35.

80. UTERGA (P. Fr. Esteban de). *Nociones elementales del idioma Goagiro con su correspondiente vocabulario*. Rome, 1895.

81. VERNEAU (R.) et RIVET (P.). *Ethnographie ancienne de l'Équateur*. Mission du Service géographique de l'armée pour la mesure d'un arc de méridien équatorial en Amérique du Sud, 1899-1906. Paris, t. VI, fasc. 1, 1912.

82. WHIFFEN (Thomas). *The North-West Amazons ; notes of some months spent among cannibal tribes*. Londres, 1915.

L'ÉCRITURE EN AMÉRIQUE.

Les indigènes de l'Amérique ont ignoré l'écriture véritable, à l'exception de quelques peuples de l'Amérique centrale.

Les *kipu* (*quipu*) du Pérou, assemblages de cordelettes de différentes couleurs, portant des nœuds en nombre variable, dont l'usage a persisté jusqu'à nos jours, n'étaient que des moyens mnémotechniques, destinés à rappeler, non pas certains événements, comme on l'a dit souvent, mais certains nombres se rapportant à ces événements (4) ; ces imparfaits moyens de numération ne peuvent pas plus être comparés à un procédé d'écriture que les bâtons à entailles des Dakota, des Hidatsa et des Šošon.

Les conquérants incasiques en introduisirent l'usage dans tout leur empire (les Araukan les appelaient *pron* et les Aymará *tšino*) ; mais en dehors même des limites de cet empire, des cordelettes à nœuds étaient employées à la manière des *kipu*, en Amérique du Sud, par les Indiens de la région de Popayán (Colombie), qui les appelaient *guappas*, et leurs voisins les Kokonuko, les Totoró et les Moguex, par les tribus des Guyanes française, hollandaise et anglaise, par les Karib de l'Orénoque, en Amérique centrale, avant l'usage des codex, par les Mexicains, qui leur donnaient le nom de *nepohualtzitzin*, et en Amérique du Nord, par les Powhatan de Virginie (famille Algonkin), les Pueblos, les Yakima (famille Šahaptin), les Paloni (sous-tribu des Tšemehuevi, famille Uto-Aztęk), et les Eskimo.

Un curieux instrument, encore en usage dans la province de Huarochirí (Pérou central), procède du même principe que le *kipu*. C'est un rôle de travail, formé d'une planchette rectangulaire, où sont inscrits les noms des ouvriers qui participent à un travail collectif de nature nettement religieuse. A côté de chaque nom, la planchette est perforée, et dans cette perforation, passent des cordons de différentes couleurs, munis de nœuds,

à l'aide desquels on note non seulement les absences et la nature du travail exécuté, mais aussi les vêtements spéciaux, le petit sac à coca, la calebasse à poudre de chaux, et les outils que chaque ouvrier doit apporter, et même le zèle plus ou moins grand dont il fait preuve dans l'accomplissement de sa tâche (**6**, 534).

Les nombreuses pictographies, que les Indiens du Nouveau-Monde dessinaient sur leur propre corps, la pierre, l'os, le bois, le cuir, la coquille, etc..., dont ils ornaient des bandes en perles de coquille ou *wampum*, sont des inscriptions uniquement figuratives, comparables aux « histoires sans paroles » de nos journaux illustrés. Ce qui caractérise ces pictographies, c'est qu'elles sont indépendantes de la langue de celui qui les trace et peuvent théoriquement être comprises par tous. Elles avaient atteint leur plus haut degré de développement chez les Kiowa et les Dakota.

Quant aux pétroglyphes, si nombreux dans les deux Amériques, ils n'avaient probablement un sens précis que pour celui qui les exécutait et pour les membres du petit groupe auquel il appartenait. Autant qu'on peut s'en rendre compte, ces signes se rapportaient bien plus souvent à des faits personnels au dessinateur, qu'à l'histoire de la tribu dont il faisait partie. Dans certaines régions, il semble cependant qu'ils étaient en rapport avec certains mythes ou certaines pratiques religieuses (**2**). En tous cas, il est illusoire de chercher à les interpréter, comme on l'a tenté trop souvent.

En fait, on ne trouve une écriture proprement dite que dans l'Amérique centrale.

Les Nahua ou Aztẹk se servaient de peaux de cerf et, plus souvent, d'une espèce de feutre mince en fibres d'agave, recouvert d'un enduit calcaire, où ils dessinaient et peignaient des hiéroglyphes. Ces manuscrits ou *codex* sont parvenus jusqu'à nous en assez grand nombre, souvent avec commentaires ou traductions en Espagnol ou en Nahuatl transcrit en caractères latins.

L'écriture nahuatl fait encore une très large part aux figures purement descriptives, qui rappellent les pictographies dont nous venons de parler; mais elle emploie, en outre, des signes phoné-

tiques. Certaines pictographies, représentant des objets sous une forme plus ou moins stylisée, ont pris en effet une valeur phonétique et ont été combinées entre elles pour constituer les diverses syllabes de mots, sans que le sens de ces mots ait le moindre rapport avec le sens des signes composants. On faisait ainsi des rébus véritables, souvent très mauvais, car les différents sons ainsi associés ne rendaient qu'approximativement le son du mot composé. Par exemple, le nom de lieu *tollantzinco* était représenté par un faisceau de roseaux, *tollan*, et la partie postérieure d'un corps humain, *tzin-tli*, le locatif *-co* restant sous-entendu.

Le scribe avait d'ailleurs le droit de choisir entre plusieurs signes phonétiques homophones pour rendre la même syllabe ; ainsi, la syllabe *quauh* pouvait être figurée par un arbre, *quauh-itl*, ou un aigle, *quauh-tli*.

On voit par ces exemples combien le système était imparfait; cette imperfection apparaît surtout lorsque les indigènes ont cherché à transcrire les noms propres espagnols. Par exemple, le nom du vice-roi Antonio de Mendoza est figuré par un agave, *me-tl*, et un rongeur indigène, *tozan*, soit : *me-tozan*.

En résumé, les *codex* nahua renferment des éléments figuratifs, qui théoriquement peuvent être interprétés sans connaissance du Nahuatl, et des éléments phonétiques, pour le déchiffrement desquels cette connaissance est indispensable. C'est exactement le cas des rébus, dont certaines parties peuvent être lues dans n'importe quelle langue, mais dont certaines autres n'ont de sens que dans une langue déterminée (**1**, 345-364 ; **3**).

On connaît également des codex, présentant de grandes similitudes avec les manuscrits nahua, exécutés sous l'influence de l'ancienne civilisation zapotęk (manuscrits xicalanques), d'autres qui sont l'œuvre directe des Zapotęk et des Mixtęk. Les Kuikatęk, les Mazatęk, les Mixe-Zoke, les Tšinantęk, les Otomí, etc... en ont aussi fabriqués de semblables à une époque historique.

L'écriture maya se compose, elle aussi, de signes hiéroglyphiques, qui sont toujours régulièrement disposés en lignes verticales ou horizontales. On possède sur elle de nombreux documents : ce sont trois *codex* fabriqués de la même façon que

les *codex* mexicains sur papier d'agave, et surtout les inscriptions sculptées dont les Maya ont couvert leurs monuments. Malheureusement, en l'absence de tout texte bilingue ou simplement annoté, le déchiffrement de ces documents est à peine ébauché.

Grâce aux précieuses indications recueillies par Landa près des indigènes, peu de temps après la découverte, on est parvenu cependant à identifier les signes qui désignent les nombres, les jours et les mois. La fréquence de ces signes et leur alternance régulière dans tous les documents en notre possession montrent que ces documents se rapportent surtout à la computation du temps. Les recherches multiples, souvent ingrates, dont ils ont été l'objet, ont permis de se faire une idée assez exacte du système de calendrier des anciens Maya, mais jusqu'ici toute tentative de déchiffrement des autres caractères a échoué. Certains savants ont supposé que, comme l'écriture nahuatl, l'écriture maya était mi-figurative, mi-phonétique ; pour d'autres, elle serait purement phonétique, tandis que, pour les américanistes allemands, elle serait plutôt idéographique. Jusqu'au jour où un heureux hasard nous mettra en possession d'un document bilingue, il est peu probable que l'on puisse sortir du domaine de l'hypothèse.

On trouvera dans l'excellent travail de Morley sur les inscriptions de Copan la bibliographie complète de la question (5).

INDEX BIBLIOGRAPHIQUE.

1. BEUCHAT (Henri). *Manuel d'archéologie américaine*. Paris, A. Picard, 1912.

2. *Handbook of american Indians North of Mexico* (edited by Frederick Webb HODGE). Bureau of american ethnology, Bulletin 30, Part 2. Washington, 1910. Article *Pictographs*, p. 242.

3. LEHMANN (Walter). *Ergebnisse und Aufgaben der mexikanistischen Forschung*. Archiv für Anthropologie. Braunschweig, neue Folge, t. VI, 1907, p. 113-168.

4. LOCKE (L. Leland). *The ancient quipu or peruvian knot record*. New York, The american Museum of natural history, 1923.

5. MORLEY (Sylvanus Griswold). *The inscriptions at Copan*. Carnegie Institution of Washington, Publication nº 219. Washington, 1920.

6. TELLO (Julio C.) et MIRANDA (Próspero). *Wallallo ; ceremonias gentilicas realizadas en la región cisandina del Perú central (distrito arqueológico de Casta)*. Inca. Lima, t. I, 1923, p. 475-549.

P. RIVET.

INDEX

DES

NOMS DE LANGUES ET D'ÉCRITURES

Les noms d'écritures sont suivis de [é].

Certains noms insérés sont des dénominations géographiques ou historiques qui servent à désigner ou à situer des langages ; ils sont suivis de mots entiers, tels que [lac], ou des abréviations suivantes : [p] pays, [v] ville, village, [î] île, [r] rivière, [pop] population, tribu.

Des noms de langage suivis d'un autre nom entre crochets, mais non suivis d'un numéro de page

1° sont une variante du nom mis entre crochets, variante non insérée dans le livre ou insérée seulement dans un titre de la bibliographie,

ou bien

2° désignent un dialecte de la langue dont le nom est entre crochets, dialecte dont il n'est pas question dans le texte.

Toutes les variantes et tous les noms de dialectes indiqués dans le texte figurent à l'index à leur ordre alphabétique.

L'ordre suivi est l'ordre français ; les caractères à signe diacritique sont entremêlés avec les caractères simples, ainsi *ś* avec *s*, *ḥ* et *ḫ* avec *h*, etc.

Il n'est pas tenu compte de ' et ʿ.

Note. — Certaines variantes, notamment des noms usuels en anglais, ayant pu être négligées, tenir compte des recommandations suivantes :

un nom qui ne se trouve pas sous C doit être cherché sous K et inversement ;

un nom qui ne se trouve pas sous CH doit être cherché sous TCH, *tš* ;

un nom qui ne se trouve pas sous J doit être cherché sous DJ, *dž* ;

un nom qui ne se trouve pas sous SH, doit être cherché sous CH, *š* ;

et ainsi de suite pour d'autres variations orthographiques possibles.

a

h (ḥ, ḫ)

l

p (p')

v (v')

w (ẅ)

ADDITIONS ET CORRECTIONS[1]

Page 18. Sur la linguistique générale, consulter plusieurs livres récents :

E. CASSIRER, *Philosophie der symbolischen Formen*. I. *Die Sprache*. Berlin 1923.

Festschrift STREITBERG, *Stand und Aufgaben der Sprachwissenschaft*. Heidelberg 1924 (sur les langues indo-européennes).

H. DELACROIX, *Le langage et la pensée*. Paris 1924.

Holger PEDERSEN, *Sprogvidenskaben i det nittende aarhundrede. Metoder og resultater*. Copenhague 1924.

O. JESPERSEN, *The philosophy of grammar*. Londres, 1924.

Page 21, ajouter à la note 2 : A. Cuny, *Études prégrammaticales sur le domaine des langues indo-européennes et chamito-sémitiques*. Paris, 1924.

P. 43, sur la carte, lire *Kurdistan*.

P. 57, ajouter à la note 5 : K. Nyrop, *Grammaire historique de la langue française*, 4 volumes parus, Copenhague (Paris), 1899-1924.

P. 73, haut. — Sur une autre variété d'anglais exotique, dite Beach-la-Mar ou Sandalwood english, voir Jespersen, *Language*, p. 216.

— ajouter à la note 3 : G. Gerullis, *Die altpreussischen Ortsnamen*. Berlin-Leipzig, 1922.

— ajouter à la note 4 : *Litauische Mundarten*, gesammelt von

1. Quelques indications figurant ici ne sont pas signalées à l'index.

A. Baranowski, herausgegeben von Fr. Specht. Leipzig, 1920.

P. 77, ajouter à la note 2 : Franz Tetzner, *Die Slawen in Deutschland*. Braunschweig, 1902.

P. 84, l. 14, voir addition à la p. 21.

P. 88 et 130. — *ś* ou *ś* de l'égyptien est une sifflante non déterminée (*s* mouillé ?).

P. 90, l. 5 du bas, au lieu de *dŏbarö-xa*, lire *dŏbaröxa*.

P. 93, haut. — La partie *cultivable* de la Mésopotamie ne devait pas excéder 40.000 kilomètres carrés.

P. 96, l. 2, au lieu de *aq-ṭi-rib-ma*, lire *aq-ṯi-rib-ma*.

P. 96 et 99. — La question de l'écriture sémitique doit être reprise depuis les découvertes de M. Montet à Byblos en 1923 : une inscription phénicienne remonte approximativement à 1250-1200, d'après M. R. Dussaud (voir *Syria*, 1924, fascicule 2).

P. 99, l. 7 du bas, *après* Phéniciens, *ajouter* : sémites.

P. 102, l. 9-10, rétablir : Mais déjà les derniers textes rédigés pendant cette période.

P. 112 haut. — Il faut tenir compte de l'extension de l'écriture sémitique en Asie par l'intermédiaire des Sogdiens, antérieurement à l'influence nestorienne (voir en outre p. 215).

P. 114, l. 15. — L' « ourmien » est ainsi appelé parce qu'il avoisine le lac d'Ourmia.

P. 121, l. 4 du bas, au lieu de *Ḥadramawt*, lire *Ḥaḏramaut*

P. 126, l. 8, *au lieu de* couchitiques, *lire* couchitique

P. 136, l. 9, au lieu de *tamaziġt*, lire *tamazigt*

P. 140, l. 9 du bas, au lieu de *Möṭmaṭa* de *Tamözraṭṭ*, lire *Möṯmaṯa* de *Tamözraṯṯ*.

P. 155, l. 21, ajouter : M. Szynnyei (voir Bibliographie) distingue encore dans le groupe finnois le dialecte INGRIEN parlé dans l'Ingermanland (gouvernement de Pétrograd) par 13.000 personnes environ et le LUDIQUE, parlé au Nord et à l'Ouest de Pétrosavodsk (gouvernement d'Olonets). Il ne cite pas de chiffres en ce qui concerne ce dernier parler qui est en voie d'extinction.

P. 158 l. 16, ajouter : M. Kai Donner (voir Bibliographie), propose de remplacer la dénomination traditionnelle de samoyède

kamassique par celle de samoyède du Sud ou samoyède du Sayan.

P. 158, l. 21 du haut, ajouter : Ce chapitre était déjà imprimé quand, à la suite de recherches entreprises depuis plusieurs mois, il nous a paru possible d'émettre l'hypothèse d'après laquelle l'eskimo serait apparenté aux langues ouraliennes qui ont fait l'objet de notre étude dans le présent volume. Il est encore trop tôt pour se prononcer sur la nature exacte du lien de parenté qui les unit (voir Eskimo, p. 611, avec renvois bibliographiques).

P. 161, l. 4, au lieu de *vĕkon'* « mine », lire *vĕkon'* « mince ».

P. 162, l. 18, au lieu de *sunan* « traîneaux », lire *sunàn-* « traîneaux ».

— l. 19, au lieu de *nämän* « noms », lire *nämän-* « noms ».

P. 163, l. 21, *au lieu de* un datif en **-k*, *lire* un latif en **-k*.

P. 165, l. 6, au lieu de : *inessif pluriel*, lire *illatif pluriel*.

P. 169, l. 19, au lieu de *kåpyŏk* « ils prennent », lire *kåpyäk* « ils prennent ».

P. 182 bas, ajouter E. Lewy, *Tscheremissische Grammatik*. Leipzig, 1922.

P. 183, l. 2, au lieu de *Ulber*... lire *Ueber*...

P. 223, l. 9-10, *au lieu de* Bargourinsk *lire* Bargouzinsk.

P. 224, l. 1, *au lieu de* Konkou-nor *lire* Koukou-nor

P. 237, l. 9, *au lieu de* King-tcheou *lire* King-Tcheou.

P. 261, ajouter : A. Eckardt, *Koreanische Konversations-Grammatik*. Heidelberg, 1923.

P. 272, ajouter : Waldemar Bogoras, *Chukchee*, dans : Handbook of american indian languages, by Franz Boas. Bureau of American Ethnology, Bulletin 40, part 2. Washington, 1922, p. 631-903.

P. 313, avant **Mitannien**, ajouter : A. Poebel, *Grundzüge der sumerischen Grammatik*. Rostock (chez l'auteur), 1924.

P. 315, avant **Hétéen**, ajouter. C. Franck, *die Altelamischen Steininschriften*, Berlin, 1923.

P. 317 haut, ajouter : L. Mayer and J. Garstang, *Index of Hittite names*. British School of Jerusalem, 1923.

Anatolian Studies presented to Sir William Ramsay. Manchester, 1923.

P. 318 bas, voir Addition à la page 21.

P. 368, l. 11, *au lieu de* mohongiā *lire* mohongiă.

— l. 23, *au lieu de* lushēī, *lire* lushĕi.

P. 373, l. 17, *au lieu de* où, *lire* ou.

P. 374, l. 8 du bas, ajouter : B. Karlgren, *Analytic dictionary of chinese and sino-japanese*. Paris, 1923.

P. 396, l. 20, *au lieu de* Muöng, *lire* Muong.

P. 398, l. 17, *au lieu de* lumineūx, *lire* lumineux.

P. 421, l. 12 du bas, *au lieu de* kamber, *lire* kamberi.

P. 428, l. 3, *au lieu de* babuhan, *lire* labuhan.

P. 440, l. 1, *au lieu de* adiṇ, *lire* adiṅ.

P. 561, l. 7, ajouter : en s'appuyant sur un mémoire de Froberville paru dans le *Bulletin de la Société d'histoire naturelle* de l'Ile Maurice d'avril 1846.

P. 573, l. 19, au lieu de *mmaze*, lire *ni maze*

P. 579, l. 1 — *ou* — ou

— l. 7 — *Nyaro* — *Nyoro*

— l. 19 — KÌRUNDI — RUNDI

— l. 20 — *rundi* — *kìrundi*

P. 581, l. 14 — NYAMWZI — NYAMWEZI

P. 587, l. 13 — *Bengua* — *Benga*

P. 597, l. 2 du bas, *au lieu de* p. 605, *lire* p. 627

P. 630, l. 2, *au lieu de* dialectes *lire* dialecte.

P. 636, l. 4 du bas, *au lieu de* Tšinantęk, *lire* Tšinantęk.

P. 647, l. 9 du bas, *au lieu de* Matši ganga *lire* Matšsiganga.

P. 682, l. 5 du bas, *au lieu de* Tšibitša, *lire* *Tšibtša.

P. 712. — Pendant le tirage de ce livre, trois importants mémoires ont été publiés sur les langues américaines :

Radin (Paul), *The relationship of Maya to Zoque-Huave*. Journal de la Société des Américanistes de Paris, nouv. série, t. XVI, 1924 — Dans ce travail, l'auteur donne une série remarquable de concordances lexicographiques et morphologiques entre le Maya et les langues du groupe Mixe-Zoke.

Aza (Fr.-José Pío), *Estudio sobre la lengua machiguenga*. Lima,

Casa editora « La opinión nacional », 1924 — Excellente étude grammaticale du dialecte Matšiganga, qui appartient au groupe pré-andin de la famille Arawak.

Mata (Fr.-Pedro de la), *Arte de la lengua Cholona*. Inca, revista trimestral de estudios antropológicos. Lima, t. I, n° 3, juillet-septembre 1923, p. 690-750 — Grammaire intéressante de la langue Tšolona, sur laquelle on ne possédait que des documents insuffisants.

P. 724, col. 724, après ligne 2, ajouter : Ba-tua 595.

— l. 4, *lire* : Baudocitarae 684.

— l. 2 du bas, supprimer : Bendiapa 665.

P. 747. A l'index, on ne trouve pas le mot *japhétique*.

M. Marr admet qu'il faut distinguer un petit groupe de langues encore aujourd'hui représenté par les langues caucasiques et le basque, et dont auraient fait partie plusieurs langues mortes telles que le sumérien, l'élamite, les anciennes langues de l'Asie antérieure, l'étrusque. C'est ce groupe de langues qu'il nomme japhétique. Les langues indo-européennes auraient en notable partie remplacé ces langues, et se seraient mélangées avec elles.

Il a été fondé à Leningrad (Petrograd) un *Institut de recherches japhétiques*, qui a déjà publié deux fascicules d'un *Jafetičeskij sbornik* (Recueil japhétique), Peterburg 1922, Petrograd 1923.

CARTES ET CROQUIS DANS LE TEXTE

Pages.

PLANCHES HORS TEXTE

TABLE DES MATIÈRES

MACON, PROTAT FRÈRES, IMPRIMEURS.